AD-A043 572 HONEYWELL RADIATION CENTER LEXINGTON MASS F/G 17/5
AUTOMATED IMAGE ENHANCEMENT TECHNIQUES FOR SECOND GENERATION FL--ETC(U)
AUG 77 DAAG53-76-C-0195
UNCLASSIFIED HRC-77-5-1 NL

1 OF 9
AD A043572

(1)

(6) AUTOMATED IMAGE ENHANCEMENT
TECHNIQUES FOR
SECOND GENERATION FLIR.

(9) INTERIM REPORT.

Submitted to
HONEYWELL SYSTEMS AND RESEARCH CENTER

(11) AUGUST 1977 (12) 169p.

(14) HRC-
Document No. 77-5-1

(15) DAAG53-76-C-0195

HONEYWELL
Radiation Center
2 Forbes Road
Lexington, Massachusetts 02173

404 486 JB

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

SECTION 1

SUMMARY

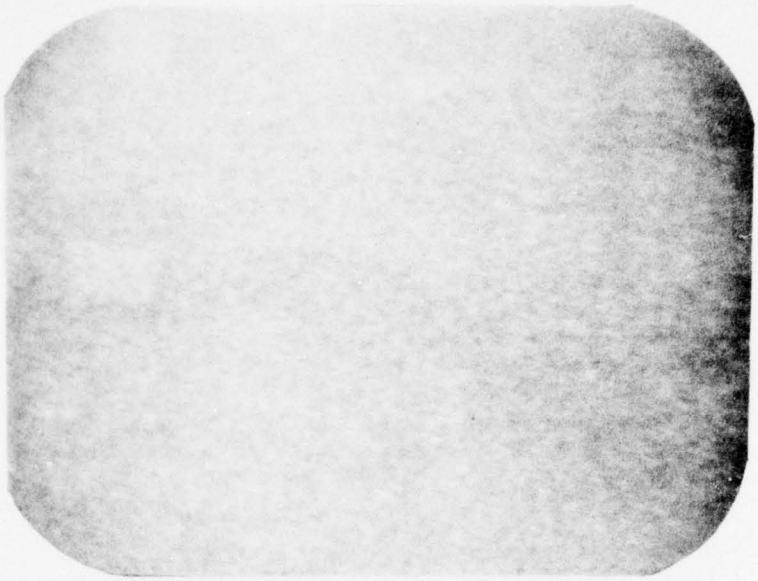
This document reports on the work done by Honeywell Radiation Center for the Honeywell Systems and Research Center under Phase I of their contract DAAG53-76-C-0195 with the U.S. Army Night Vision Laboratory.

Section 2 covers the primary image processing functions (in a second generation FLIR) for the derivation of a video signal, free of artifacts, from the detector array. To illustrate the signal processing functions (multiplexing, time-delay-and-integration, anti-blooming, background cancellation, gain equalization) a representative second generation array of 750 x 3 elements is assumed. Relevant aspects of the CCD technology (e.q. dynamic range) associated with these processing functions are included.

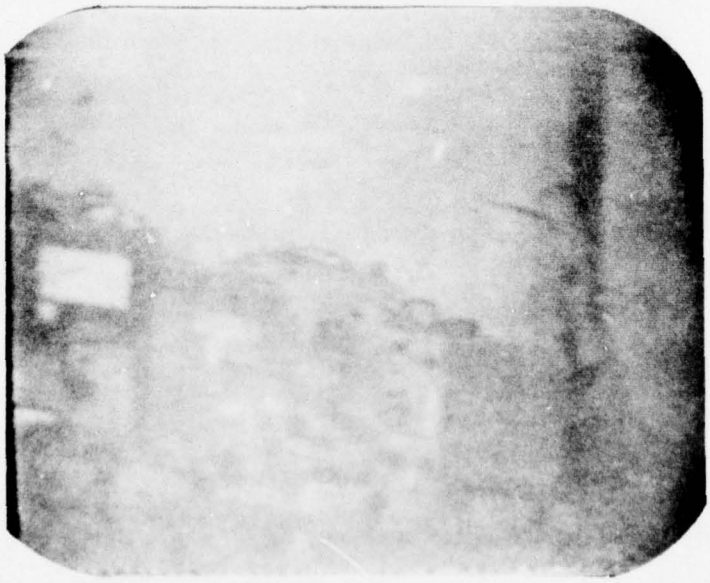
Section 3 deals with interframe processing. It is shown how successive frames of moving FLIR video may be put into registration--making possible any of a number of useful interframe processing functions. For example, integration for improved signal-to-noise ratio (i.e., improved thermal sensitivity) as illustrated in Figure 1.1. (A 256 x 256 pixel, 60 fps real-time, moving image integrator is being constructed in 1977 at the Honeywell Radiation Center on internal funds.)

Auto-focus is covered in Section 4. The various factors which can cause a FLIR to be defocused are discussed. In particular the (operator's) "focusing difficulty," in correcting defocus due to variations in object range, is quantified. Three widely different focus-correcting techniques are described and it is shown how their applicability, in any particular case, depends upon the basic cause of defocus in each case.

In Section 5 two approaches to the problem of contrast enchancement are presented.



a) NOISY LLLTV IMAGE (STANDARD 525-LINE, 30 FPS)



01396

b) AFTER 2 SECONDS REAL-TIME INTEGRATION IN DIGITAL MEMORY (256 x 256 PIXEL)

Figure 1.1 SIGNAL-TO-NOISE IMPROVEMENT THROUGH REAL-TIME INTEGRATION

SECTION 2

PRIMARY PROCESSING FUNCTIONS

First generation FLIRs utilize relatively small numbers of detectors. For a number of reasons, including, for example, the need for higher sensitivity and a higher field-of-view/resolution ratio, a large number of detectors must be used on the focal plane in a second generation FLIR. A representative application might call for 2250 detectors arranged in a 750 x 3 format. A one dimensional parallel scan would generate a 750 x 1000 element picture at 30 frames per second (Figure 2.1). The three elements in the scan direction would be used for a TDI (time delay and integration) function, to improve sensitivity and allow adequate responsivity to be maintained in the event that a few isolated detector elements were defective.

To obtain useful video information from this focal plane, certain signal processing functions must be implemented. These include signal multiplexing, antiblooming, and gain equalization. In addition, depending on the system design, background cancellation may be necessary. These signal processing functions, termed primary processing functions, will give a video output from the focal plane which is free of spurious image artifacts and has a minimum noise level. This output will be suitable as an input to the more advanced signal processors which could include contrast enhancement, resolution restoration, and others.

The division of the primary processing functions between on-focal-plane and off-focal-plane locations is mainly dictated by the state-of-the-art of focal plane technology. Presently, the use of a charge coupled device (CCD) is the only feasible way to implement several of the primary processing functions. Therefore, the discussion of the implementation of the functions given below will be directed towards designs which use CCDs.

2.1 FOCAL-PLANE PROCESSING

In principle, each of the 750 detector signal lines could be directed to its own LED (also on the focal plane) via individual (low frequency) gain, offset and antiblooming circuits. The LED would be viewed by reflection in the scan mirror. However, the mechanical and electrical complexity of such an implementation preclude its use in a practical system.

To avoid the necessity for 750 individual signal lines from the focal plane, the detector outputs must be multiplexed into a single video signal. The multiplexing function may be realized by using a hybrid PV detector/silicon CCD focal plane structure. A typical configuration is shown in Figure 2.2, which shows a staggered linear array with two stages of TDI.

The detector array is mounted adjacent to, or on the silicon CCD. Interconnects are fabricated by photolithographic technology.

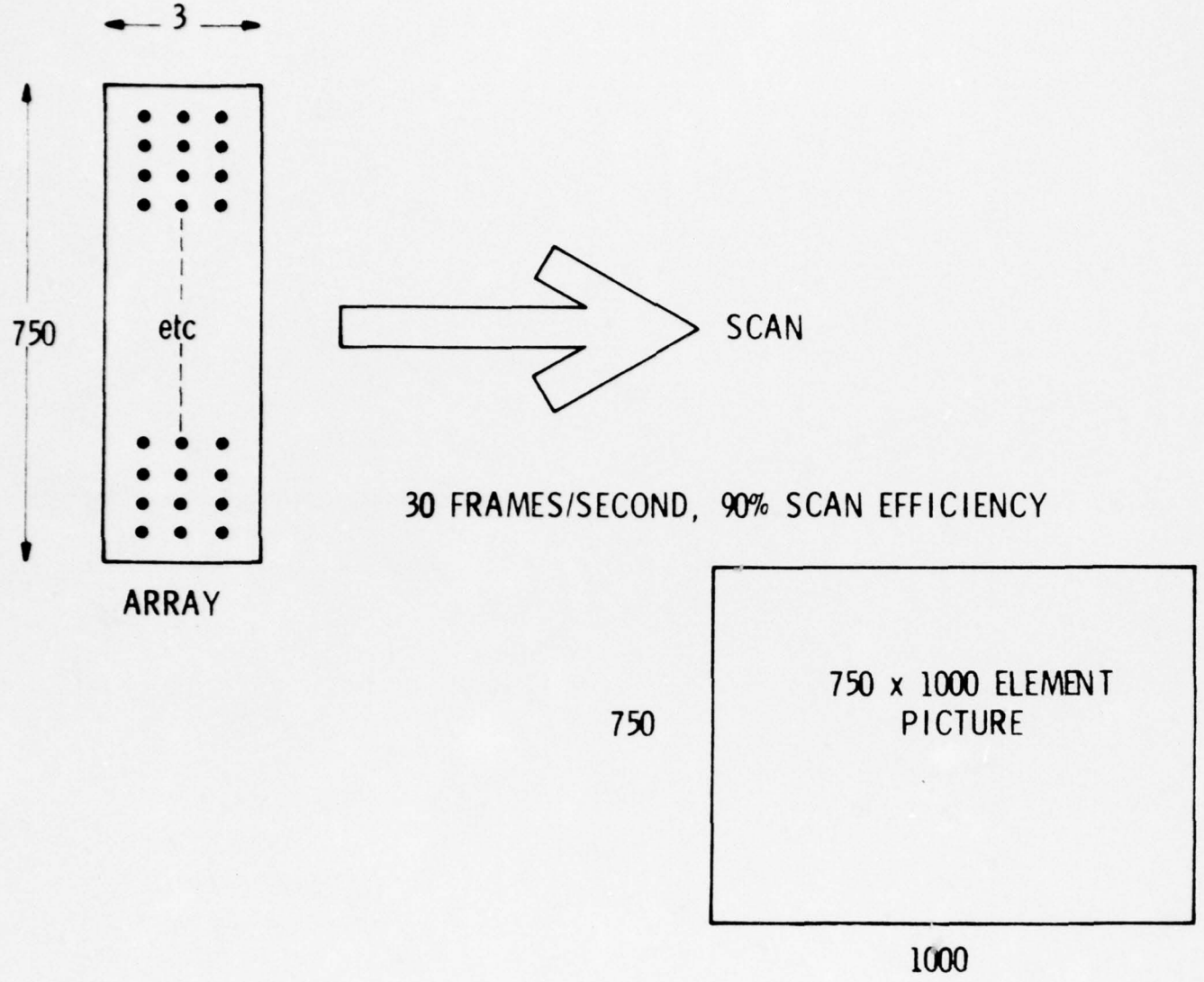


Figure 2.1 ILLUSTRATIVE ARRAY SCAN SYSTEM

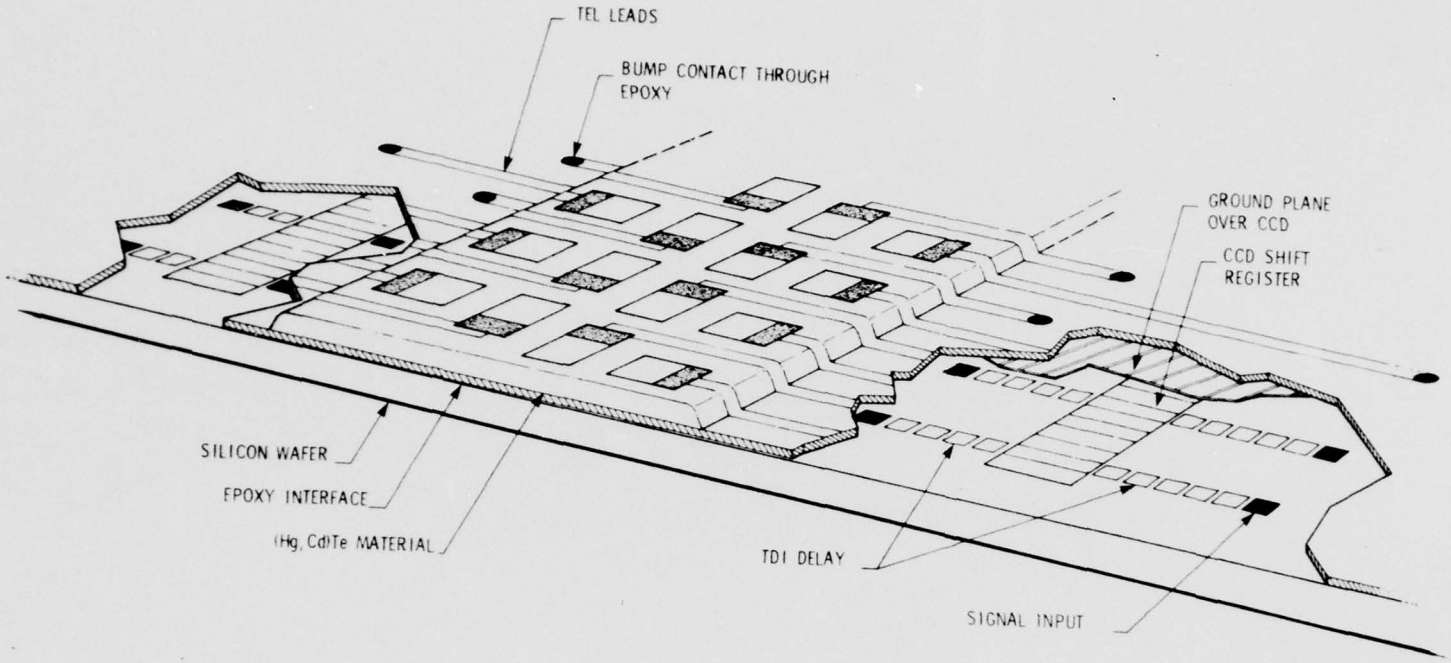


Figure 2.2 HYBRID FOCAL PLANE STRUCTURE

2.1.1 Detector Coupling

The detectors are to be multiplexed by parallel injection of detector signals into a CCD shift register. To maintain the detector S/N performance, the CCD input, transfer, and output electronics must not introduce any excess noise. The CCD noise is predominantly due to the input of the detector signal into the CCD storage well; therefore, an analysis of the input circuit electronics is sufficient to predict the performance of the entire detector/CCD circuit. Depending on the specific system operational requirements, either ac or dc coupling of the detectors to the CCD may be used.

For background limited (BLIP) photovoltaic detectors, direct coupling of zero biased detectors will give no noise peformance degradation (Reference 2.1). In this case, all of the detector current due to signal and background flows into the CCD.

For non-BLIP photovoltaic detectors, the signal may be directly coupled to the CCD with no excess input noise if the detectors are operated with a moderate reverse bias, in the neighborhood of 50 mV (Reference 2.1). Then, the CCD input current consists of signal and background current along with the detector saturation current.

Under certain conditions of spectral sensitivity band, bias, and integration time, it may be difficult to accommodate all of the charge integrated in the storage well in the CCD shift register. For that case, if S/N considerations discussed below allow it, a background subtraction circuit may be used to reduce the required charge handling capacity of the CCD. This circuit may be implemented by special structures on the focal plane. The applicability of a background subtraction circuit is discussed in Section 2.1.2.

For systems using both BLIP and non-BLIP detectors, ac coupling could relieve the problem of too high CCD input current and the associated short integration times without being limited by the background subtraction capability. However, there are serious problems with ac coupling at the CCD input for advanced FLIR applications discussed below, which limit its usefulness.

2.1.1.1 S/N and Dynamic Range Considerations for Direct Coupled BLIP Detectors

Assuming that the detectors operate with BLIP performance, and that no excess noise is introduced by the CCD input circuit, the required integration well capacity may be directly related to the detector NETD and operating wavelength.

The integrated input charge in the CCD is due to the background, and is given by:

$$Q = q\eta\Phi_B A_d \tau_I$$

The signal charge due to a target ΔT is given by ΔQ, where:

$$\Delta Q = q\eta\Phi_B A_d \tau_I A(\lambda_c, T_o)\Delta T$$

where $A(\lambda_c, T_o)$ is the contrast function. Values of $A(\lambda_c, T_o)$ are given in Table 2.1 for selected values of λ_c and T_o (background temperature).

Table 2.1

$\lambda_1 - \lambda_2$	$A(\lambda_1 - \lambda_2, T_o)$		
	T_o = 250 K	T_o = 300 K	T_o = 350 K
3 - 4.2 μm	0.058	0.041	0.029
3 - 5.5 μm	0.052	0.037	0.022
8 - 14 μm	0.021	0.015	0.0075

Because the detector is operating at BLIP conditions, and there is no excess CCD noise, the charge noise (ΔQ) in the CCD input well is the statistical fluctuation in the number of electrons in the well and is, therefore, equal to the square root of the charge. Letting ΔT_N equal the target temperature change at which the S/N is one, we have:

$$\Delta Q_N = \sqrt{Q_m} = \Delta Q$$

Then, the minimum storage well size for the desired T_N is given by:

$$Q_m = \frac{1}{A^2(\lambda_c, T_o)\ \Delta T_N^{\ 2}}$$

Using values $A(\lambda_c, T_o)$ given in Table 2.1, we find the minimum integration well capacities shown in Figure 2.3 for the several wavelength bands and background temperatures of interest in order to obtain the required NEΔT of ΔT_N.

The maximum signal-to-noise (S/N), or dynamic range, for the minimum well size and for no background subtraction is given simply by:

$$DR = \frac{Q_m}{\sqrt{Q_m}} = \frac{1}{A(\lambda_c, T_o)\Delta T_N}$$

The system with the smallest T_N has the largest dynamic range, and (from Figure 2.3) also required the largest CCD well size.

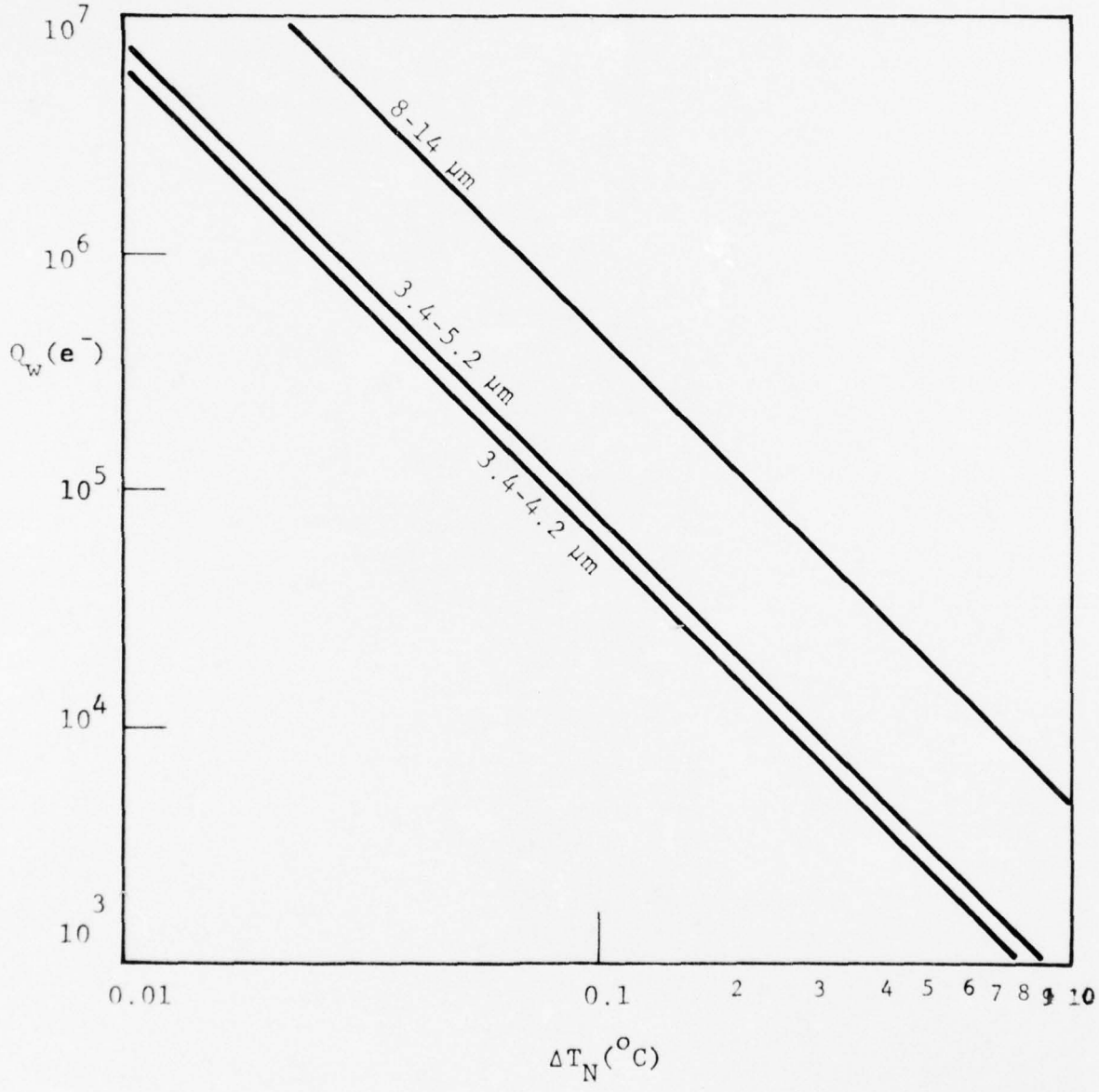


Figure 2.3 MINIMUM WELL CAPACITY FOR A GIVEN ΔT_N AT T_Q = 300 K (BLIP DETECTORS, DIRECT COUPLED; MULTIPLY BY $(1 \mp \alpha/2)^2$ FOR NON-BLIP DETECTORS WHERE $\alpha = 2\ I_{sat}/I_B$).

77-5-1

2.1.1.2 S/N and Dynamic Range Considerations for Non-BLIP Direct Coupled Detectors

For non-BLIP detectors operated at reverse bias to optimize detector/CCD S/N, the total CCD input current is:

$$I = I_s + I_B + I_{sat}$$

The total integrated charge, in the absence of signal, is:

$$Q = (I_B + I_{sat})\tau_I = q\eta\phi_B A_d \tau_I + I_{sat}\tau_I$$

The magnitude of I_{sat} with respect to I_B determines how close to BLIP the detector will operate.

A convenient measure of the closeness to BLIP operation of the detector is the ratio α of twice the saturation current to the background current:

$$\alpha = \frac{2I_{sat}}{I_B}$$

For ideal PV detectors, $I_{sat} = \frac{kT}{qR_o}$, so we have:

$$\alpha = \frac{2kT}{q^2\eta\phi_B R_o A_d}$$

Then, for the detector alone, operated at zero bias,

$$D^* = \frac{D^*(BLIP)}{(1+\alpha)^{1/2}}$$

The total integrated CCD charge from a reverse biased PV detector is then given by:

$$Q = q\eta\phi_B A_d \tau_I \left(1 + \frac{\alpha}{2}\right)$$

For non-BLIP detectors $\alpha > 1$ and the CCD charge for reverse bias detectors is greater than that for BLIP detectors under the same background conditions.

The CCD input circuit still introduces no excess noise, so the noise in the integration well is still equal to $\sqrt{Q}$, as for BLIP detectors.

However, the signal charge for a given ΔT is a smaller part of Q than obtained from BLIP detectors. The required minimum well size for a given ΔT_N is derived from setting ΔQ for a minimum signal equal to the noise in the well, ΔQ_N:

giving $$Q_m = \left(\frac{1 + \frac{\alpha}{2}}{A(\lambda_c, T_o)\Delta T_N} \right)^2$$

For a detector with an $R_o A_d$ such that the detector D* is half BLIP, Q_m is over six times that required for a BLIP detector.

The dynamic range for the non-BLIP detector with the minimum CCD well is given by:

$$DR = \frac{Q_m}{\sqrt{Q_m}} = \frac{1 + \frac{\alpha}{2}}{A(\lambda_c, T_o)\Delta T_N}$$

2.1.1.3 AC Coupled Detector/CCD Inputs

The use of ac coupling between the detector and the CCD can reduce charge handling requirements on the CCD under certain conditions. The principal condition is that signal gain must be used between the detector and the CCD. A typical input configuration is shown in Figure 2.4. Because of variations in device properties following the coupling capacitor from input to input, there are gain as well as offset variations which must be corrected. In addition, the amplifier itself introduces additional gain variations. Therefore, there is no fundamental advantage in the use of ac coupling to reduce drastically the gain and offset correction requirements.

The factor which precludes the use of ac coupling for advanced FLIR applications is that the preamplifier and coupling capacitor dissipate power and take up too much area on the focal plane. For future applications which may have over 750 detector channels with numerous detectors in a TDI configuration for each channel, it will be impossible to allocate sufficient power to the focal plane cooler, and to obtain the necessary high detector packing density.

2.1.2 Background Subtraction

The detector output (charge) is dependent on the absolute temperature of the scene element that is imaged onto the detector and on the amount of detector bias. The useful scene-structure information is generally the temperature-modulation (contrast) in the scene. Consequently, the dc component of the detector signal (corresponding to the minimum scene temperature) contains no useful information.

To the extent that this "background" dc signal may limit dynamic range, it should be removed. The removal of this background signal directly at the focal plane may be necessary to keep the signal within the dynamic range of the focal plane CCD, but it can introduce additional artifacts which must then be subsequently corrected. If the background component is removed off-the-focal-plane, after the detector signals have been multiplexed into a single video signal, no such artifacts are introduced.

Background subtraction can provide increased dynamic range for a given well size, or reduce the required CCD size. The amount of background

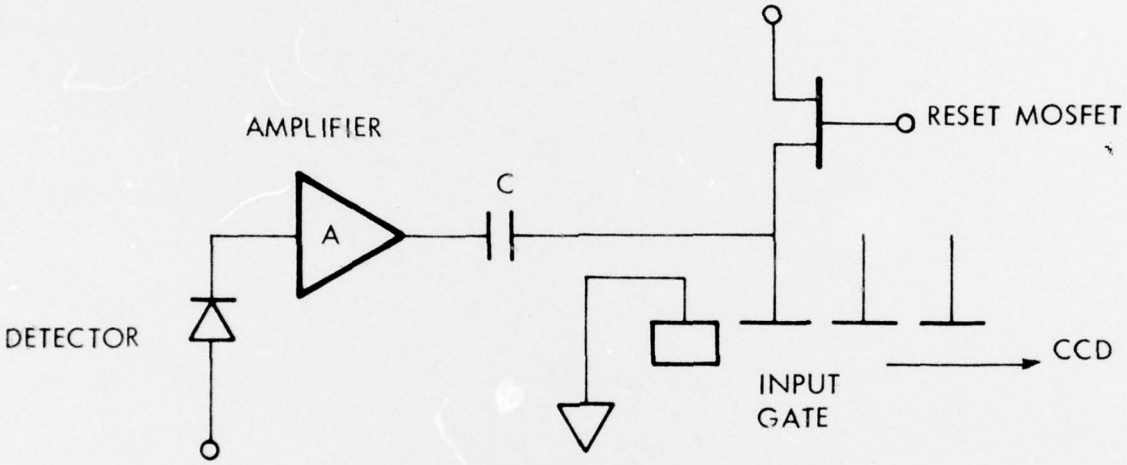


Figure 2.4 TYPICAL AC COUPLED DETECTOR/CCD INPUT CIRCUIT

77-5-1

subtraction which may be used depends both on S/N considerations and detector uniformity considerations. A S/N limitation on the amount of background subtraction for BLIP detectors which is possible for the different wavelength bands is derived here. Assuming that the target temperature is ΔT_s above the background, we have the maximum charge due to the signal, Q', given by:

$$Q' = Q + \Delta Q = Q \left[1 + A(\lambda_c, T_o) \Delta T_s \right]$$

For optimum background subtraction, we retain and inject into the CCD only the charge ΔQ due to the target ΔT_s. The ΔQ, or CCD capacity with background subtraction, depends on the maximum target temperature T_s, the noise equivalent temperature T_N, and the spectral band in a way given by:

$$\Delta Q = A(\lambda_c, T_o) Q_m \Delta T_s = \frac{\Delta T_s}{A(\lambda_c, T_o)(\Delta T_N)^2}$$

This relation is plotted in Figure 2.5 for a ΔT_N of 0.01°C and T = 300 K. Also shown is Q_m, the CCD well capacity determined by noise requirements at ΔT_N with no background subtraction. For low ΔT_s, the CCD well size is decreased from that required with no background subtraction ($\Delta Q < Q_m$); while for high ΔT_s, the CCD well size is determined by the dynamic range requirement and background subtraction is of no avail.

For present detectors, detector uniformity capabilities give a second limitation on attainable background subtraction.

For BLIP detectors, the total charge injected into the CCD integrating well is the background plus signal charge. For detector responsivity variations of $\pm$ 5%, the total injected charge will be in the range given by:

$$0.95\ Q_B + 0.95\ Q_s \leq Q \leq 1.05\ Q_B + 1.05\ Q_s$$

The background subtraction circuit is designed to remove a fixed amount of charge from every input integration well, so that the CCD shift register does not have to accommodate an excessively large charge packet. In order to not subtract signal from a low responsivity detector's integration well, the subtracted charge must not exceed $0.95\ Q_B$. Subtracting this amount of charge will leave $0.1\ Q_B + 1.05\ Q_s$ in a high responsivity detector integration well, a maximum reduction of background charge of 10:1. Even this background subtraction is only of avail when the dynamic range considerations discussed above allow it.

For non-BLIP direct coupled detectors, background subtraction is even more desirable than for BLIP detectors, because the current from the detector is increased by the diode saturation current, which for non-BLIP detectors is greater than the background current.

At best, all of the diode saturation current would be subtracted, along with the scene background generated current. Detector uniformity becomes

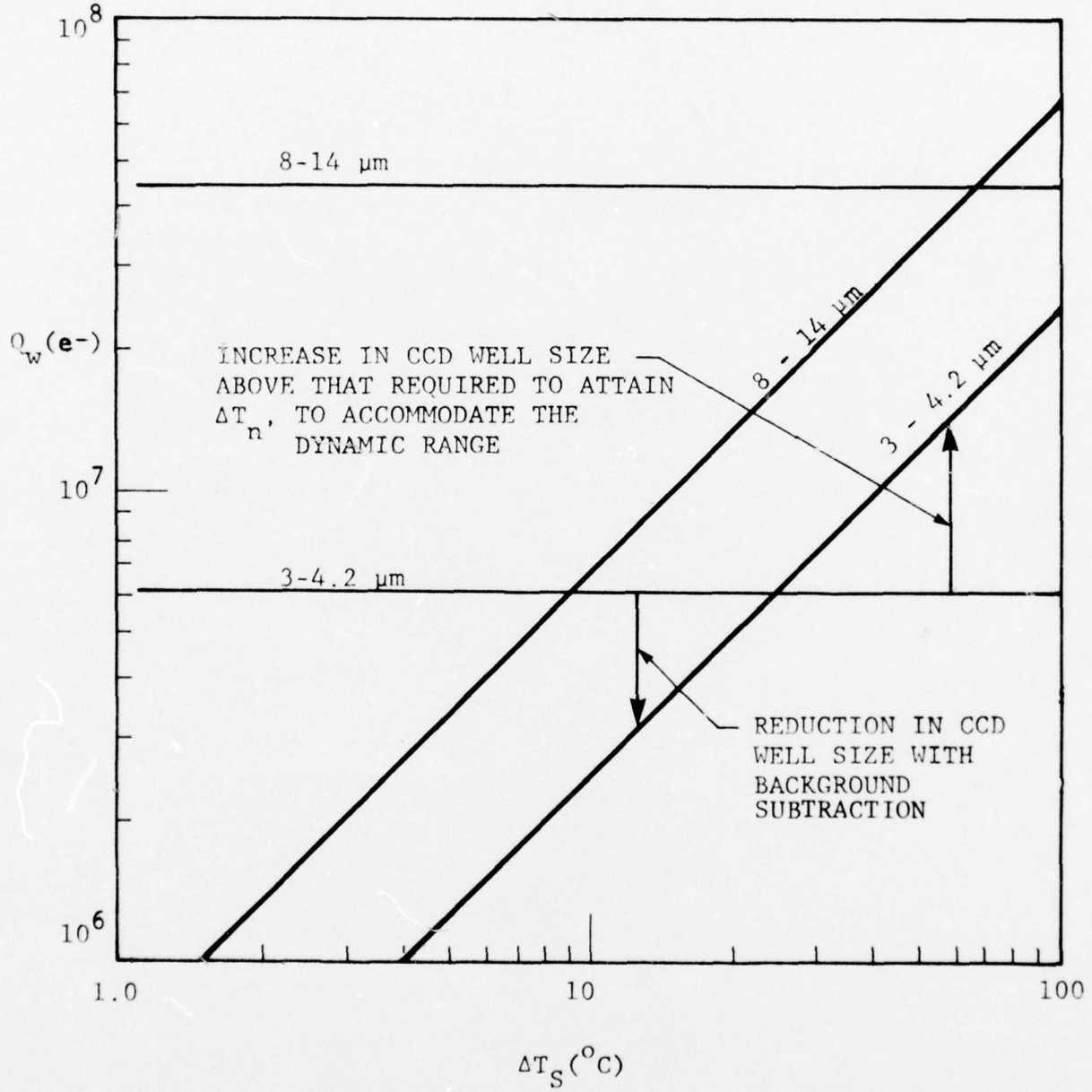


Figure 2.5 CCD WELL SIZE WITH BACKGROUND SUBTRACTION FOR $\Delta T_N = 0.01^oC$, AT $T_B = 300$ K (BLIP DETECTORS; MULTIPLY HORIZONTAL LINES BY $(1 + {}^o\alpha/2)^2$ FOR NON-BLIP DETECTORS WHERE $\alpha = 2\ I_{sat}/I_B$)

a critical parameter, because variations in detector saturation current will now limit the amount of background subtraction attainable. For example, for a $\pm$ 20% uniformity in saturation current (corresponding to about $\pm$ 20% in detector $R_o A_d$, and a $\pm$ 5% uniformity in detector responsivity, the integrated CCD charge lies in the range given by:

$$0.95\ Q_s + 0.8\ Q_{sat} + 0.95\ Q_B < Q < 1.05\ Q_s + 1.2\ Q_{sat} + 1.05\ Q_B$$

For a detector operating at one-half BLIP D*, $Q_{sat} = 3\ Q_B$, so we have:

$$0.95\ Q_s + 3.35\ Q_B \leq Q \leq 1.05\ Q_s + 4.65\ Q_B$$

Then, the most charge we can subtract from the integration well is $3.35\ Q_B$. This leaves $1.3\ Q_B$ remaining in the fullest well; i.e., there is an excess charge larger than the background generated charge remaining in the fullest well, so no net background subtraction can be achieved.

2.1.3 Antiblooming

Even though the dynamic range of CCD detector combination may be designed to cover the desired range of scene temperatures, there may exist occasional "hot spots" where saturation may be inevitable; for example, in viewing hot engine exhausts. For these situations, antiblooming circuitry must be included in the focal plane in order to localize the effects of a saturation picture element to that element.

This antiblooming function may be readily incorporated into the hybrid detector/CCD structures discussed here. A schematic layout of an antiblooming circuit is shown in Figure 2.6a. The detector is connected either directly to the source S (dc coupled), or through an amplifier to the integrating gate G_I (ac coupled). Charge is integrated into the storage well W, and then transferred by gate G_T into the CCD. The section G_{A-B}, D is the antiblooming circuit, whose operation may be understood by reference to the dogleg section through A-A' shown in Figure 2.6b. If a hot spot induces too much current flow into the storage well W, the excess will flow into the antiblooming drain D rather than into the CCD, because G_{AB} is set below G_T. This circuit will eliminate blooming for signals up to about 100 times the maximum signal within the dynamic range. The hot spot will appear only as one saturated spot on the output, and will not affect any outputs from adjacent or other detectors.

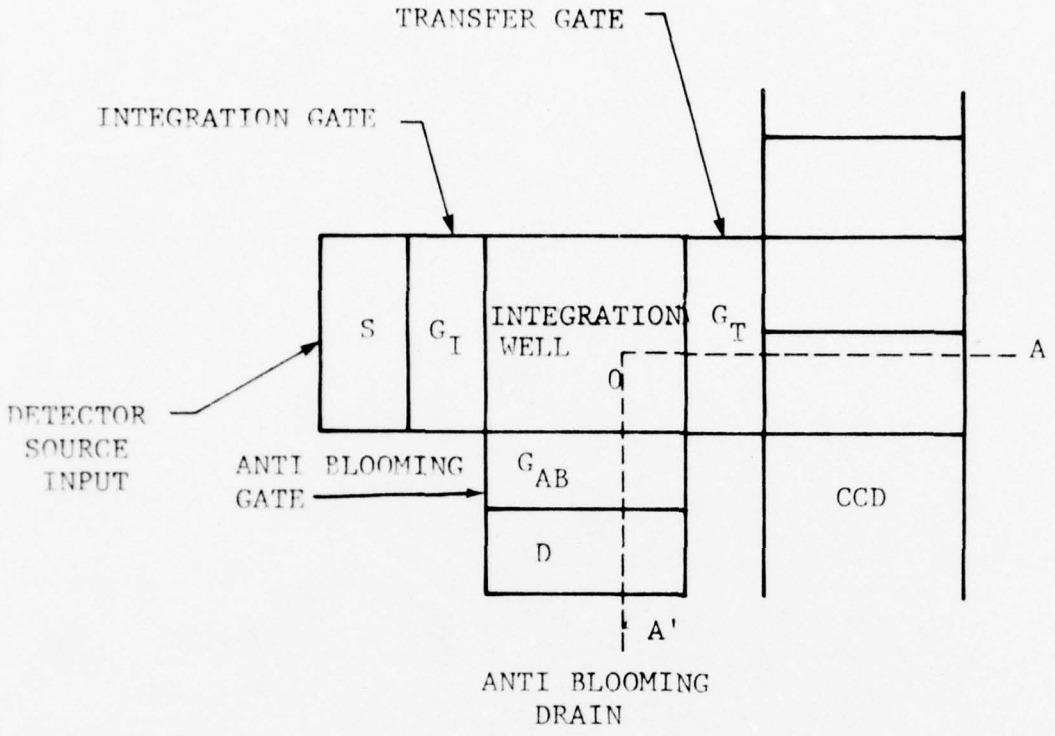


Figure 2.6a *ANTIBLOOMING CIRCUIT SCHEMATIC*

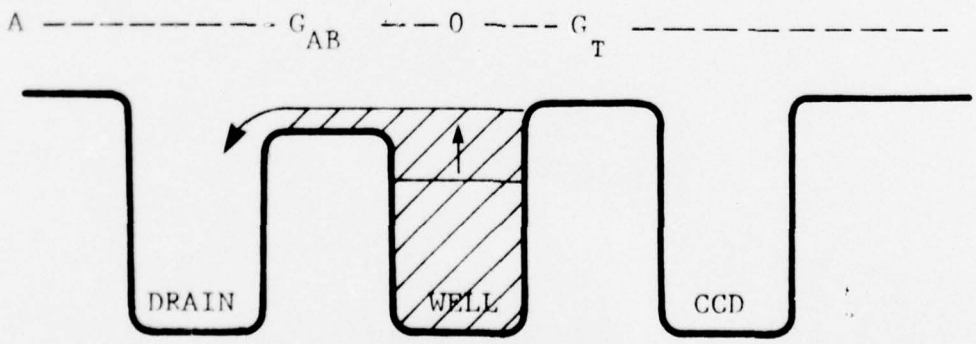


Figure 2.6b ANTIBLOOMING CIRCUIT OPERATION - SECTION
SECTION A-A' OF FIGURE 2.6a

2.1.4 Time Delay and Integration (TDI)

The time delay and integration function is simply implemented using a CCD. The individual detectors are connected to adjacent inputs on the CCD. The shift register rate is adjusted to be in synchronism with the scan rate of the image across the detector array. Then, detector signals from the same image points are summed N times into one moving CCD well. The CCD well capacity must be adjusted to accommodate N times the individual detector charge output. An improvement of $\sqrt{N}$ in S/N has been demonstrated using a CCD implemented TDI focal plane.

2.1.5 CCD Geometrical Considerations

The minimum integration well and CCD sizes (Figures 2.3 and 2.4) depend on the spectral sensitivity bands, the required $T_N(NE\Delta T)$, and the dynamic range requirements. The maximum practical well size is determined by real estate availability on the focal plane and charge transfer dynamics. Practical surface channel integration wells can hold 1×10^{12} charges/cm^2, and buried channel CCDs can hold half of that number. Then, reasonable sizes of integration wells ($\sim 1 \times 10^{-5}$cm^2) can hold enough charge to satisfy minimum requirements (1×10^7 electrons). The more critical consideration is how long can the detector current be integrated into the CCD. The integration time to obtain 0.8 of full well is given, for a BLIP detector by

$$\tau_I = \frac{0.8\, Q_W}{q\, \eta\, \Phi_B A_d}$$

Where Q_W is the charge in the well, Φ_B is the incident radiant flux, A_d is the area of the detector, η the detector quantum efficiency and q the electronic charge. For a well size of 10^{-5}cm^2, a 0.001-inch square detector with a quantum efficiency of 0.75 and an f/1 optical system, the maximum integration time is given in Table 2.2. The integration times are shorter for non BLIP, reverse biased detectors by a factor of $(1 + \alpha/2)^{-1}$.

Table 2.2

$\lambda_1 - \lambda_2$	τ_I
3 - 4.2 μm	3.2 ms
3 - 5.0 μm	0.64 ms
8 - 14 μm	8.7 μs

For the 8 to 14- μm band, the integration time is short; so, high speed CCD would be required for readout of a large detector array. Larger input well sizes could make the 8 to 14- μm system marginally satisfactory. The short wavelength systems have no problems with integration time.

2.1.6 Summary

In summary, direct coupling of PV detectors to the CCD is suitable for focal plane applications so long as wavelength band and integration times are compatible with reasonable CCD and integration well sizes. Excellent noise performance may be obtained. Background subtraction is of marginal utility, if large ranges of background temperatures or large dynamic ranges are necessary, or if large variations in detector responsivity and/or saturation current exist. Ac coupling is not practical for future FLIR systems requiring large numbers of detectors on the focal plane.

2.2 OFF-FOCAL-PLANE PROCESSING

2.2.1 Gain and Offset Equalization

The remaining primary processing functions are gain equalization and level equalization. That is, if the intensity (temperature) of the ij^{th} scene element is T_{ij}, the corresponding video level is V_{ij} where:

$$V_{ij} = a_j T_{ij} + b_j$$

After ideal gain and level equalization:

$$V_{ij} = aT_{ij} + b$$

The gain and level equalization functions could, in principle, be performed on the focal plane. However, with the current state of the focal plane technology, it is recommended that these functions be performed off-the-focal-plane. Their implementation off-the-focal-plane is relatively straightforward and simplifies the focal plane design.

The requirements for offset and gain compensation are illustrated in Figures 2.7 a and b, which show how the CCD input charge might vary from one detector to another for BLIP detectors (Figure 2.7a) and non-BLIP detectors with reverse bias (Figure 2.7b).

The choice of which correction (gain or offset) is performed first and the amount of correction required depend on whether or not the detectors are BLIP. The two possibilities illustrated in Figure 2.7 will be discussed in detail.

2.2.1.1 BLIP PV Detectors Direct-Coupled to the CCD

With direct coupled BLIP PV detectors (Figure 2.7a), the CCD input charge is given by:

$$Q = (I_B + I_s)\tau_I = q\eta\phi_B A_d \left[1 + A(\lambda_c, T_o)\Delta T_s\right] \tau_I$$

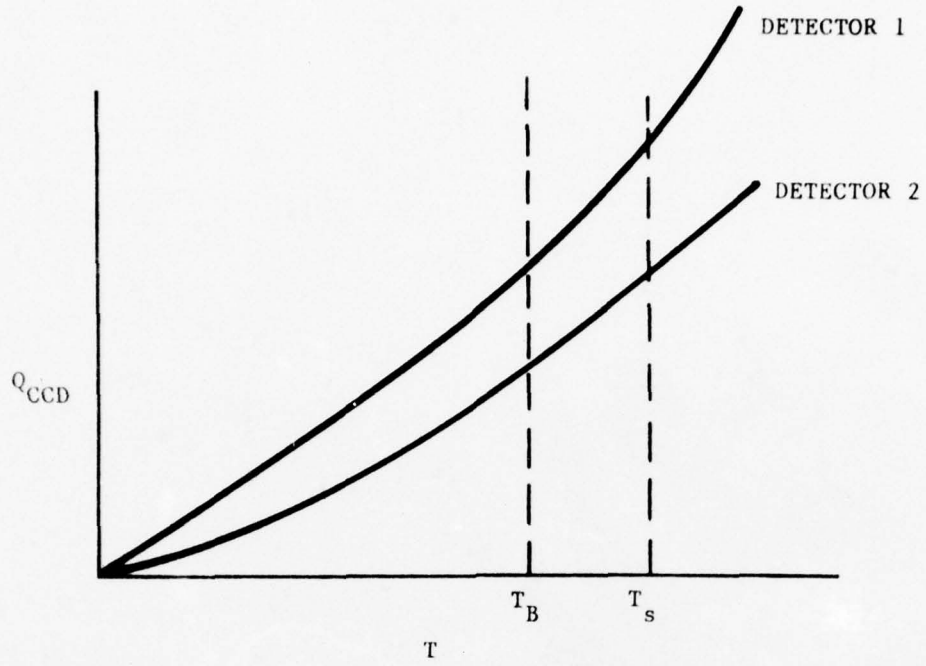


Figure 2.7a CCD INPUT CHARGE VARIATIONS WITH BLIP DETECTORS

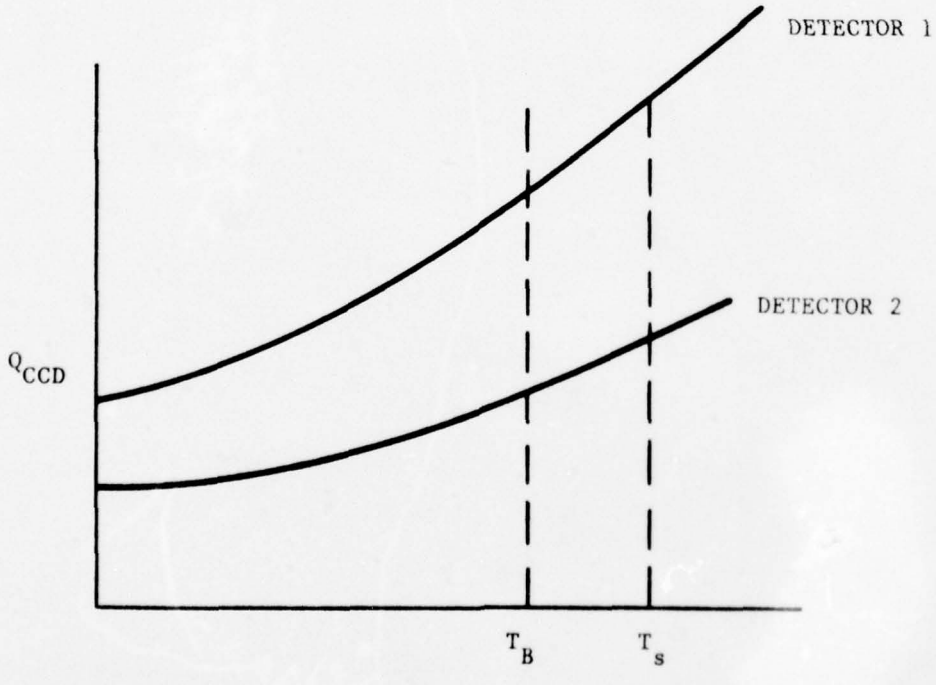


Figure 2.7b CCD INPUT CHARGE VARIATIONS WITH REVERSE BIASED NON-BLIP DETECTORS

In this case, the nonuniformities of the detectors are manifested in the detector responsivity variations, which enter in the above relation in the product $\eta\phi_B A_d$ and in $A(\lambda_c, T_o)$.

The first step in correction is to look at a cool uniform field and normalize the individual detector outputs for offset (actually due to the effect of responsivity variations on the background). This normalization takes account of responsivity variations, which may range up to $\pm$ 10%. The precision of the correction must be such that the normalized outputs are equal to within about half of the NETD. The precision is expressed by the ratio of 10% of the background charge to the change due to one-half of the NETD:

$$\text{Precision} = 2\left\{\frac{0.1\ q\eta\phi_B A_d \tau_I}{0.5\ q\eta\phi_B A_d \tau_I A(\lambda_c, T_o)\Delta T_N}\right\} = \frac{0.4}{A(\lambda_c, T_o)\Delta T_N}$$

(The factor 2 comes from the necessity of handling responsivities above and below the mean responsivity).

From Table 2.1, for a 3 to 5.5-μm detector and a ΔT_N of 0.05 K, the precision required is 216, or 8 bits in a digital correction system. The normalization may be done either by multiplication or addition.

The second correction is made when the array scans over a second uniform field at a temperature T_2 a small temperature above the background. The precision required for this gain correction depends on whether multiplication or subtraction was used for the first correction. Multiplicative correction will remove the effect of responsivity variations on the background charge, and then the gain dynamic range correction is reduced to correcting the variations:

$$\Delta Q = \tau_I q\eta\phi_B A_d\ \Delta A(\lambda_c, T_o)\Delta T_s$$

to half the noise $\Delta Q_N = \tau_I q\eta\phi_B A_d A(\lambda_c, T_o)\Delta T_N$

This requirement is:

$$\text{Precision (2)} = 2\left\{\frac{\Delta A(\lambda_c, T_o)\Delta T_s}{0.5\ A(\lambda_c, T_o)\Delta T_N}\right\}$$

The ratio of $\Delta A/A$ depends mainly on the variation $\Delta\lambda_c$ of cutoff wavelength among the detectors. Using the information of Table 2.1 and reasonable ranges of $\Delta\lambda_c$, we estimate $\Delta A/A \leq \pm 0.05$ for the 3.4-4.2 spectral band. Then, for a maximum ΔT_s of 50°C and a ΔT_N of 0.05 K, the requirement is:

Precision (2) $\approx$ 200 = 8 bits

In practice, the precision of correction may be somewhat reduced, as the minimum ΔT_N is required mainly when targets very close to the background temperature are being displayed. Then ΔT_s should be replaced by the difference between the actual background and the first reference temperature. This replacement can reduce the precision (2) to about 5 bits. Another way of looking at the reduced correction requirement is that the display can only accommodate about 5 bits of dynamic range; so, with hotter targets, the minimum resolved signal will be greater than ΔT_N.

2.2.1.2 Non-BLIP Reverse Biased PV Detectors Direct-Coupled to the CCD

With reverse biased PV detectors (Figure 2.7b), the CCD input charge is given by:

$$Q = \tau_I q \eta \phi_B A_d \left(1 + A(\lambda_c T_o)\Delta T_s\right) + \tau_I I_{sat}$$

Detector nonuniformities are in responsivity ($\eta\phi_B A_d$) contrast $\left[A(\lambda_c, T_o)\right]$ and saturation current (I_{sat}). The first correction is an offset correction, which subtracts the dc level due to I_{sat} and responsivity variations. For a concrete example of the precision required, consider a detector which is half BLIP, and has the nonuniformities discussed in 2.1.2. The variation in dc offset charge is then:

$$\Delta Q = \pm\ 0.65\ Q_B$$

This variation must be corrected to:

$$\frac{1}{2}\ \Delta Q_N = 0.5\ Q_B A(\lambda_c, T_o)\Delta T_N$$

Thus, the offset correction must accommodate a precision expressed by:

$$\text{Precision} = 2\left\{\frac{0.65\ Q_B}{0.5\ Q_B A(\lambda_c, T_o)\Delta T_N}\right\} \approx 1200\ (11\ \text{bits})\ (3\text{-}4.2\ \mu\text{m},\ \Delta T_N = 0.05^{\circ}\text{C})$$

Because of the random nature of the offset, the offset correction must be additive, in contrast to the case of BLIP detectors. Then, the second gain correction must correct:

$$\Delta Q = \Delta\left\{Q_B A(\lambda_c, T_o)\Delta T_s\right\} \quad \text{to } \frac{1}{2}\ \Delta Q_N$$

The variation in signal ΔQ now depends on variations in both Q_B and $A(\lambda_c, T_o)$ among detectors. The variation in responsivity of $\pm$ 10% leads to an equal variation in Q_B. This effect is in addition to the effect of variations in $A(\lambda_c, T_o)$. The net ΔQ is:

$$\Delta Q \approx \pm\ 0.15\ Q_B A(\lambda_c, T_o)\Delta T_s \text{ for } 3\text{-}4.2\ \mu\text{m}$$

Then, the second correction precision is:

$$\text{Precision} = 2\left\{\frac{0.15\ Q_B\ A(\lambda_c, T_o)\Delta T_s}{0.5\ Q_B A(\lambda_c, T_o)\Delta T_N}\right\} \approx 600\ (9\ \text{bits})$$

As noted above, the minimum ΔT_N is not usually required with large signals ΔT_s. Therefore, a reasonable reduced gain correction precision is obtained by setting $\Delta T_s/\Delta T_N$ equal to the display dynamic range of 5 bits. Then, the reduced precision is:

$$\text{Precision} = 2\left\{\frac{0.15}{0.5} \quad 0.32\right\} = 20 \text{ (5 bits)}$$

2.2.2 Gain and Offset Compensation Implementation

2.2.2.1 Gain Equalization

The effective sensitivity of the 750 detector channels will vary from one to another, due to the responsivity variations between detectors and to CCD input variations. For dc input with BLIP detectors, or non-BLIP detectors reverse biased, the principal cause of gain variations is responsivity variations among the detectors. These variations are due to quantum efficiency and area differences. Present detector arrays have demonstrated responsivity uniformities of better than $\pm$ 10%. The limit of attainable nonuniformities will be determined by photolithographic tolerances. The nonuniformities will be most pronounced for small detectors; for a one-mil detector, and a 0.5- μm photolithographic tolerance, the area contribution to nonuniformity would be 4%.

For ac coupled detectors, CCD threshold variations at the input MOSFET and input amplifier variations give an additional source of nonuniformity in input gain. For state-of-the-art CCD threshold variations, input gain variations may be a factor of two or higher. These variations are in addition to detector responsivity variations.

Individual defective elements will cause large gain variations. With a 3-element TDI, a defective element would cause a 33% drop in channel gain. It may be advisable to provide for jumper-wire connection of such large gain errors, particularly if the remaining gain errors are much smaller. In this way, the dynamic range of the gain adjustment circuit is not "wasted" on these exceptional gain errors.

Gain errors may be corrected using 750 digital words in a shift register, as illustrated in Figure 2.8. As the charge from each picture element is shifted out, generating a video level V_{ij}, a gain correction A_{ij} is applied

$$V_{ij}' = A_j V_{ij} = A_j (a_j T_{ij} + b_j)$$

such that

$$A_j a_j = a$$

i.e., $A_j = a/aj$

Then $V_{ij} = a T_{ij} + A_j b_j$

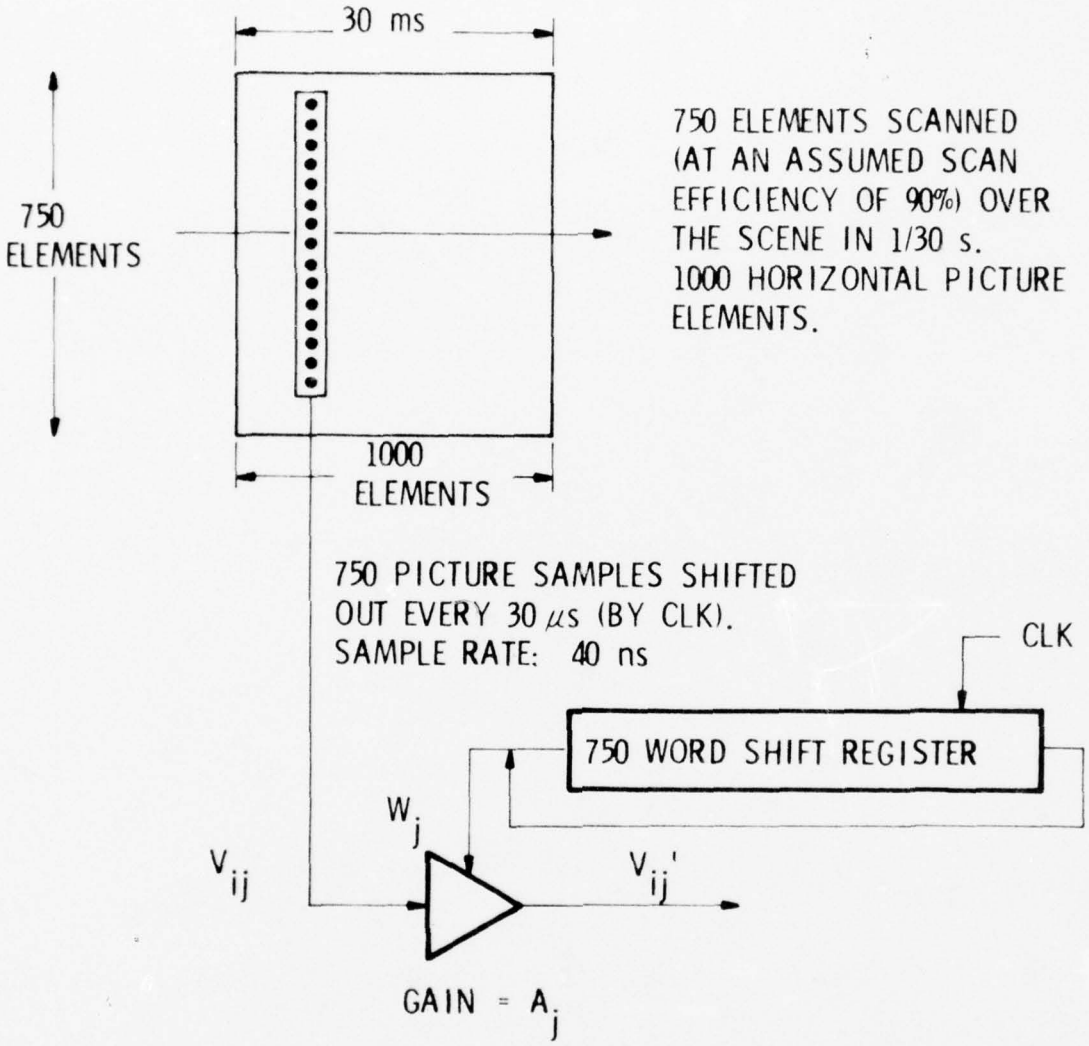


Figure 2.8 GAIN CORRECTION SYSTEM

The gain correction A_j is controlled by the j^{th} word (G_j) in the shift register. The length of the gain correction word required depends on whether or not the detectors are BLIP and on detector nonuniformities, as discussed in Section 2.1.

2.2.2.2 Offset Equalization

Even when each of the 750 individual detector channels is gain-equalized, the signal output may not be uniform when viewing a uniform temperature scene. For example, if charge subtraction is utilized on the focal plane (to improve dynamic range) the quantity of charge subtracted may not be exactly the same for each channel. For reverse biased non-BLIP detectors variations in $R_o A$ along an array will also give offset variations.

To compensate for offsets a second shift register of 750 digital words W_{ij}, as shown in Figure 2.9, is used to generate offset corrections:

$$V_{ij}' = V_{ij} - X_j = a\,T_{ij} + A_j b_j - X_j$$

By appropriate choice of X_j:

$$X_j = A_j b_j + b$$

$$V_{ij}' = a\,T_{ij} + b$$

and the effects of differing offsets (between the various channels) is removed.

2.2.2.3 Generation of the Correction Words

Four techniques for the generation of the gain and offset correction words are described.

(1) Use a two-temperature reference at the focal plane (Figure 2.10). When the array is over the T_1 temperature reference the switch S1 closes causing the gain-adjusting digital words to change according to

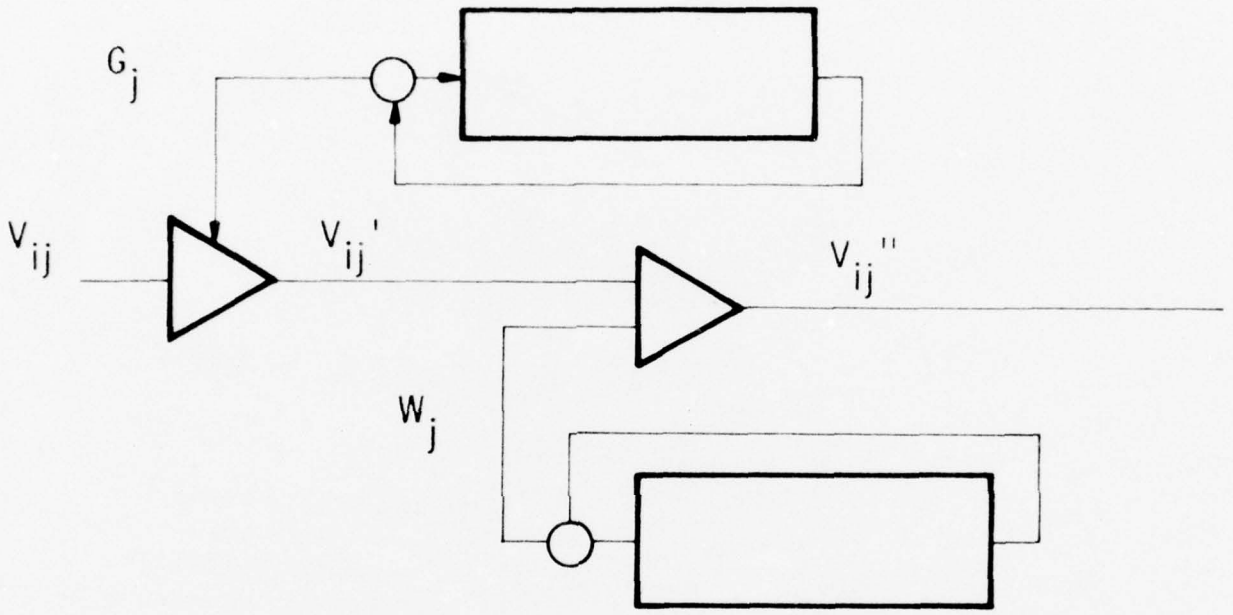
$$A_j = A_j + \lambda_1 (V - V_1)$$

This causes

$$V = A_j (a_j T_1 + b_j) - X_j \longrightarrow V_1$$

with a time constant determined by the constant λ_1. When the array crosses the T_2 temperature reference the switch S2 closes causing the offset-adjusting words to change according to

$$X_j = X_j - \lambda_2 (V - V_2)$$



$$V_{ij} = a_j T_{ij} + b_j$$

$$V_{ij}' = A_j V_{ij} = A_j (a_j T_{ij} + b_j) = a T_{ij} + b_j'$$

$$V_{ij}'' = V_{ij}' - X_j = a T_{ij} + b_j' - x_j = a T_{ij} + b$$

Figure 2.9 GAIN AND OFFSET CORRECTION

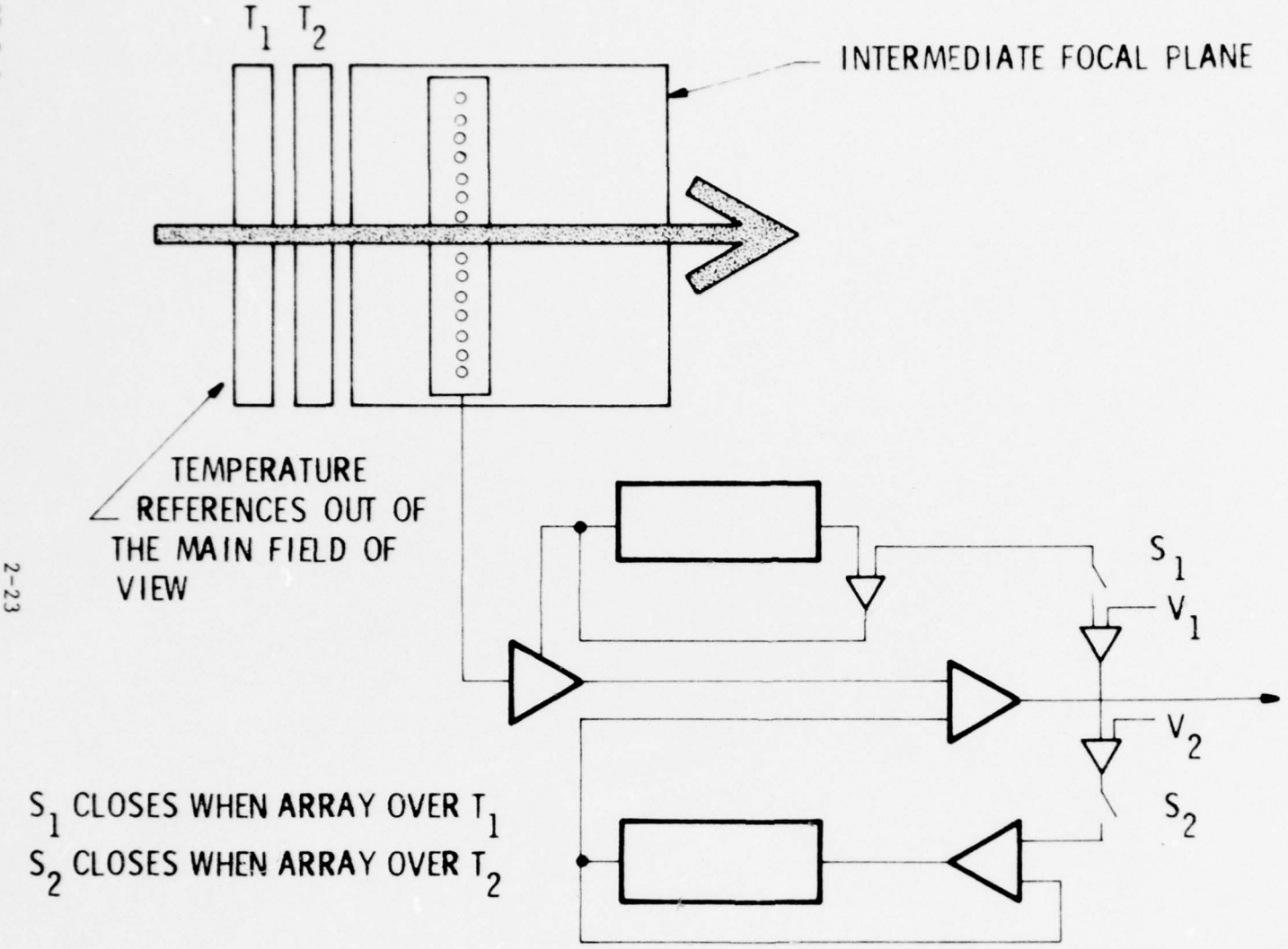


Figure 2.10 TWO-TEMPERATURE REFERENCE TECHNIQUE

This causes

$$V = A_j (a_j T_2 + b_j) - X_j \longrightarrow V_2$$

As a result of these two operations, A_j and X_j are set so that for any scene temperature T

$$V = \frac{(T - T_1)}{(T - T_2)} (V_1 - V_2) + V_1$$

(2) Precalibration (in production and/or in field maintenance) using an external temperature reference. The 1500 correction words thereby generated are stored in an EPROM (Figure 2.11). Assuming 10 bits/word a total of 15,000 bits of storage would be required. The feasibility of this technique depends upon the unknown long-term stability of the required gain and offset corrections.

(3) Use of a special self-adjust mode. In this arrangement the digital words would be generated automatically (and then stored in a RAM) each time the system is switched on. For example, by capping the system objective, or aiming it at a close-in point, a uniform (blurred) image is generated which the system could use to develop its offset correction, in a special offset adjustment mode. Gain adjustments are more difficult. They will have to be derived on the basis of equalization of the long-term variance of the signal from each of the 750 lines, as the FLIR was aimed over a range of scene images. The problem here is the time needed to effect adequate gain equalization.

(4) Continuous Correction

In this technique the gain and offset words are generated continuously, in the normal operation of the system. The principle utilized (which is described in more detail in Section 3.3.6) is that when the scene image is moving within the FLIR frame, <u>different</u> detectors are used to image the <u>same</u> scene element in successive frames.

Thus by combining successive frames, gain and offset corrections can be developed for each detector so that the video level of any particular scene element does not change as it moves through the FLIR frame (that is, as it is imaged by different detectors).

Of these four techniques. the first is relatively straightforward and can be implemented without significant development. The others have the advantage of not requiring a temperature reference but are of less certain feasibility. Further development would be required before they could be implemented.

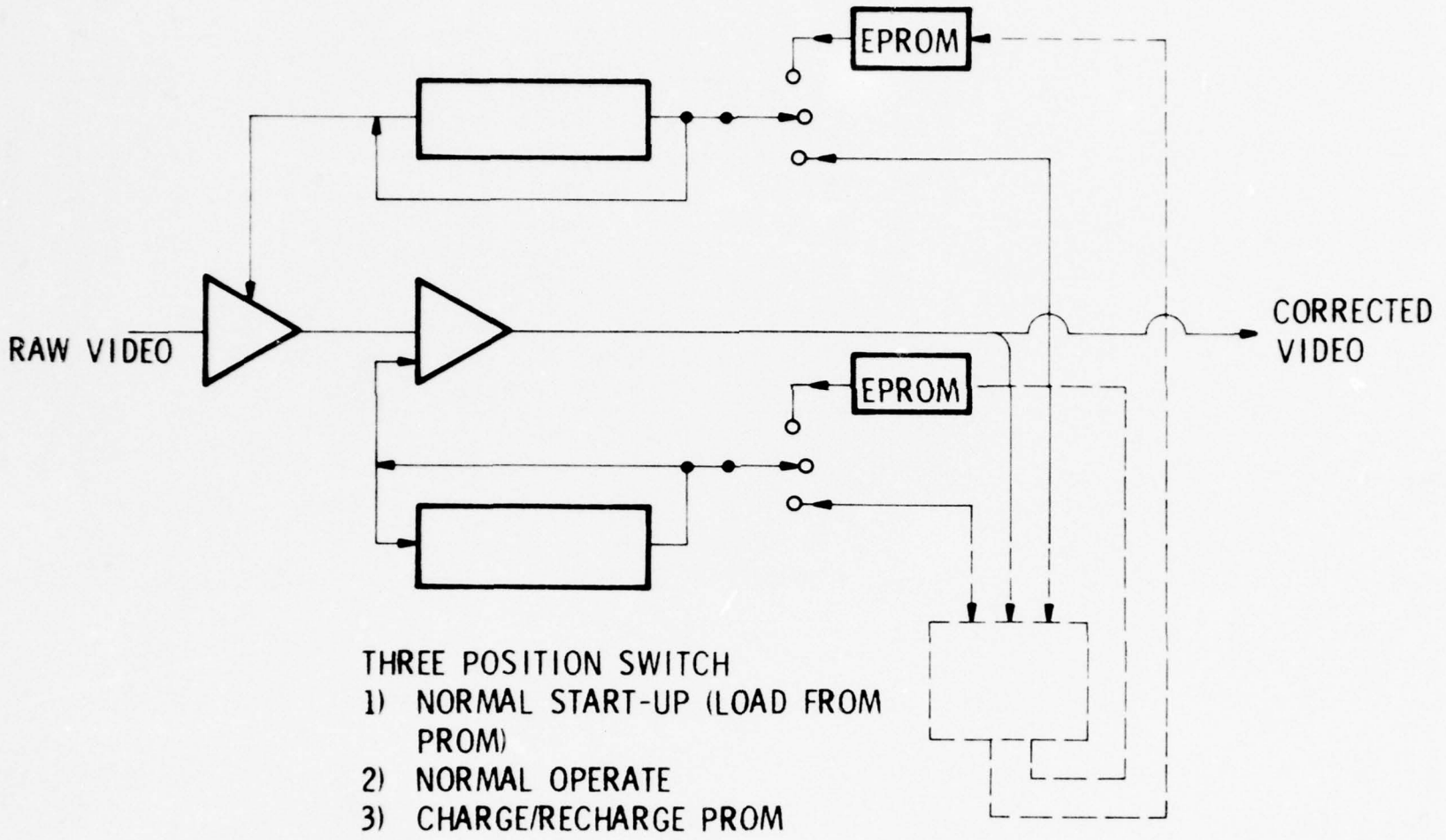


Figure 2.11 GAIN AND OFFSET CORRECTION BY EPROM

It is possible that a combination of the techniques may offer the best long term solution. For example, precalibration, with the correction words stored on a non-volatile bubble memory, together with the use of a Continuous Correction to change the contents of the bubble memory as the readout corrections change with time.

If the detector is to be ac coupled to the CCD multiplexer, then a temperature reference must be used, in the FLIR itself, to effect offset equalization. If not, the vertical MTF of the system will be seriously compromised. After the detector signals have been multiplexed into a single video line, ac coupling can be used to reduce electronic complexity. The low frequency corner determines the amount of shading that will exist across the display when viewing scenes containing large temperature differentials. The average level of the display is maintained by clamping the video (Figure 2.12) during a known output from the focal plane such as when viewing a thermal reference source.

2.2.2.4 Equalization Circuit Implementation

The key components in the equalization circuits are the memories. They must store image information from one line of detectors, which may have 1000 elements or more, and they must have adequate dynamic range, which may range up to 1000:1.

At present, these requirements are best met by digital memories, which could require 1000 words of 10 bits each, for a total of 10,000 bits. Typical power consumption for present commercial devices would be 1 watt for each memory. Access time is typically 0.2μs for a 4096 bit memory, which will limit the maximum data rate to a frequency of the order of 5 MHz. In addition, D/A and A/D converters will be necessary. Presently available commercial ten bit D/A converters can readily operate at up to 10 MHz, and dissipate about a watt.

In conclusion, it should be possible to implement the equalization and offset circuitry using digital circuits, with a data rate of 5 MHz and a dynamic range of 1000.

2.2.2.5 Temperature Reference

As indicated above, a temperature reference may be included in the FLIR to provide for a continuous gain and/or offset correction. The temperature reference(s) should be located within the FLIR optical system adjacent to a field stop. They will be just outside the active scan area (scene area) and will be viewed by the scanning array once per frame (1/30s).

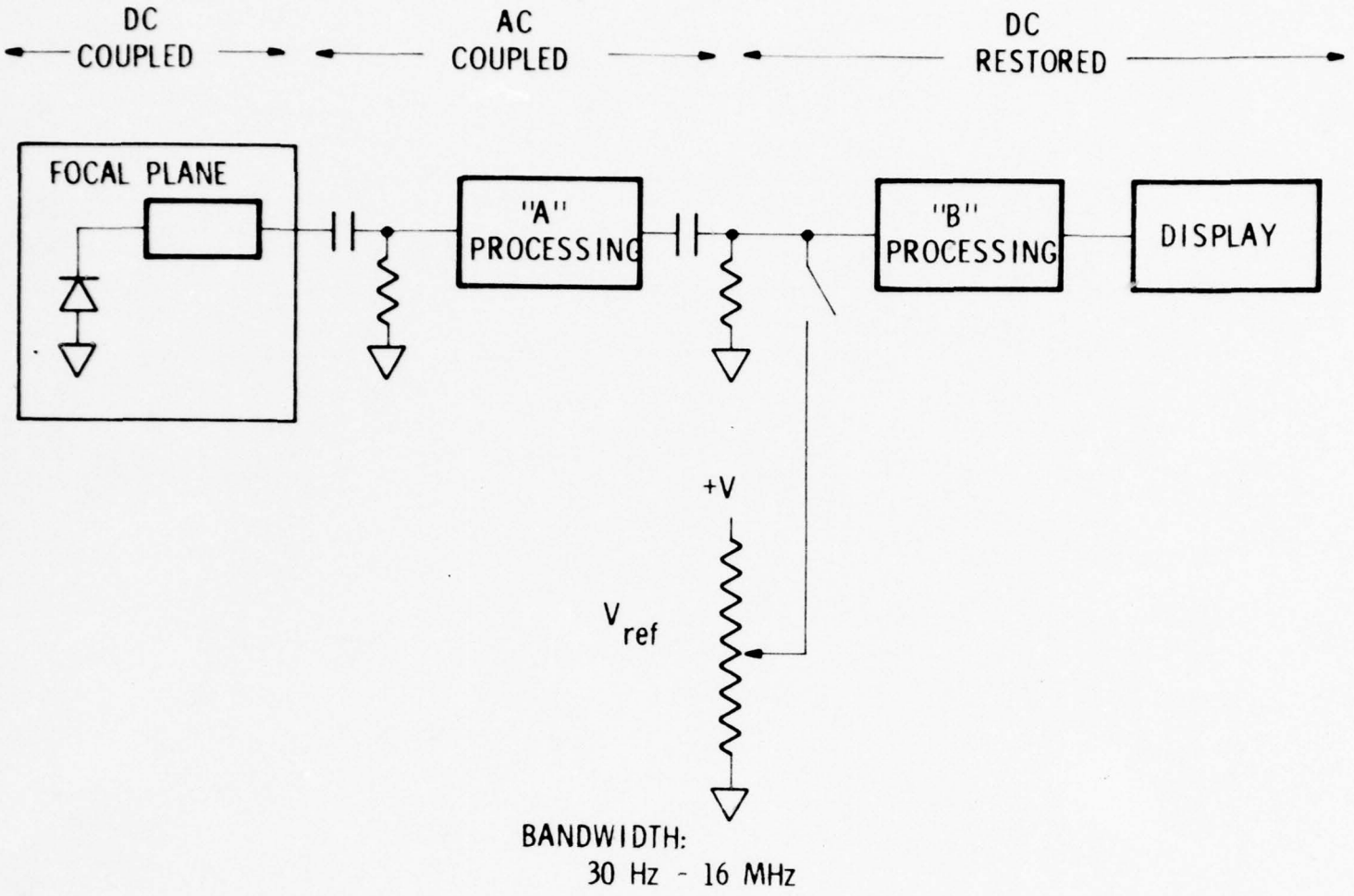


Figure 2.12 DC RESTORATION OF POST-FOCAL-PLANE AC COUPLED VIDEO

In effect, the video signals generated as the array senses the temperature reference are stored and added (or subtracted) to the video during the active portion of the scan. Thus the reference, as seen by the array, should be uniform (free of artifacts) to a greater degree than is required for the FLIR itself.

To minimize the uniformity requirements on the temperature reference, it can be located so as to be significantly out-of-focus. Then each pixel of the FLIR sees a spatial average of many (effective) pixels of the reference.

As discussed in Section 4, a FLIR will generally require a relatively few (~3) fixed-focus steps in order to cover the required focusing range. The size of the defocus blur circle increases by the order of one pixel diameter per fixed focus step.

It follows, therefore, that if the temperature reference is located at about 20 fixed focus steps away from sharp focus it will always be out-of-focus by at least 10 fixed-focus steps. That is, the defocus blur circle at the reference will be at least 10, and at most 50, pixels in diameter.

Assuming a 1 inch dimension for the field stop, the pixel dimension (at the temperature reference) will be of the order of 10^{-3} inches. To accommodate an extreme defocus condition, the active reference width should, therefore, be of the order of 0.05 inch (Figure 2.13).

The reference signal should be integrated for a number (~10) pixels (in the scan direction) so that a time average of the reference signal, rather than the unaveraged video, may be stored as reference information. If the reference were sampled for just one pixel (in the scan direction) the random noise (NETD) existing at the pixel would be "frozen" into the entire subsequent frame. It would then not only add to the system noise existing during the frame but may be more noticeable to the viewer than the system noise itself because (unlike the system noise) it would be stationary during the scan line.

The following are typical specifications* for the reference(s):

Two thermal references are required at the edge(s) of the intermediate focal plane (field stop)

*The specification will depend upon the CCD dynamic range and upon the range of scene temperatures to be accommodated.

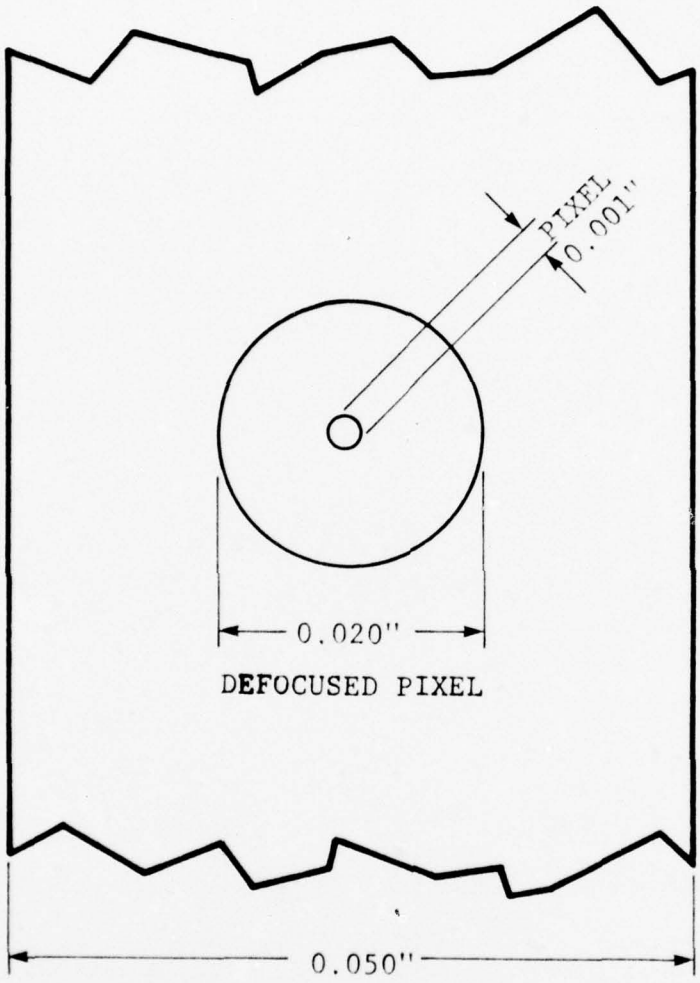


Figure 2.13 DEFOCUSED TEMPERATURE REFERENCE

They should each be approximately 1-inch long and at least 0.1 inch wide. (This guarantees that, for all focus conditions, the reference can be sufficiently out of focus that each pixel is blurred out from 0.001 inch (sharp focus) to at least 0.01 inch diameter, at the reference.)

The temperature reference should be uniform in the longitudinal direction to $0.3^{\circ}C$ for smooth end-end variations and to within 0.01° from one 0.01-inch diameter region to the next.

Case 1

One reference can be held close to case temperature. The other reference should be adjustable from $3^{\circ}C$ to $10^{\circ}C$ above the case temperature.

Case II

One reference to be adjustable from $0^{\circ}C$ to $10^{\circ}C$ below case temperature. The other from $0^{\circ}C$ to $+ 10^{\circ}C$ above case temperature.

In general, it is not necessary that the reference temperature be known accurately (e.g. $\pm 2^{\circ}C$ accuracy is adequate).

For special cases $0.1^{\circ}C$ accuracy (with $0.1^{\circ}C$ end-end uniformity) may be required.

A one minute warm-up time is desired, although a 15-minute period should also be considered.

Maintaining a thermal reference (1 inch long by 0.05 inch wide) with less than $0.3^{\circ}C$ longitudinal temperature difference can be achieved by utilizing a suitable high thermally conductive material such as OFHC copper. The reference strip of copper material would have a larger area than 1 inch x 0.05 inch specified to provide a high degree of isothermalism over the required reference area. The reference surface can be coated with a black paint such as 3M black velvet in order to have a suitably high emittance. Point-point uniformity (over 0.01 inch areas) to within $0.01^{\circ}C$ can be easily achieved.

In Case I where the reference temperature must be adjustable between 3^{o}C to 10^{o}C above the case temperature, the use of a heater bonded to the reference material with a control thermistor and suitable electronics for control can be utilized. The reference should be conductively and radiatively isolated in order to reduce gradients. Conductive isolation can be achieved with fiberglass standoffs. Radiative isolation can also be achieved with a suitable low emittance coating on all surfaces except the surface seen by the detectors.

Another approach to making a heated thermal reference is to employ a thin sheet, or rod, of a suitable electrically resistive material such as carbon. A current would be passed through the sheet or rod so as to produce internal heating. Internal heating produced in this way would be uniform except for edge effects. In this case, the length and the width and/or diameter of the strip could be larger than that viewed by the detectors. The change of resistance of the material with temperature could be used to monitor the temperature of the reference.

The other reference, which would be maintained at the case temperature, could be adequately heat sunk to the case structure.

In Case II, cooling of one reference from 0^{o}C to 10^{o}C below the case temperature (case temperature at room temperature) can be accomplished through the use of a thermoelectric (T/E) cooler. A copper reference structure would be soldered to the T/E cold platform. The T/E cold platform could be made into the same size as the reference, so that uniform cooling of the reference can be made. The power dissipation of the T/E cooler would be removed by a suitably sized finned structure. The warm reference for this case would be heated from 0^{o} C to 10^{o} C above the case temperature as was reported in Case I above. Maintenance of a controlled temperature for either reference would again be made through the use of suitable control electronics and a temperature sensor.

The reference temperature can be known to 0.1^{o}C accuracy through the use of standard commercially available thermistors.

It appears that a requirement of 0.1^{o}C end-end uniformity could also be achieved through present technology and judicious design. For this case, the thermal isolation would be higher than for Case I. Additionally, if the desired isolation could not be achieved, a photon integrating cylindrical shell of a high thermally conductive material could be utilized in order to integrate out any small gradients in the reference surface. The interior of the integrating cylinder would be a highly reflecting, diffuse surface. There would be two slits along the length in the cylinder at 90 degrees to each other. One slit would allow the energy emitted from the reference surface to enter the cylinder. Upon multiple reflections, this energy would leave the cylinder through the other slit, which would be viewed by the detectors.

It is expected that warm-up times on the order of one minute can be achieved easily. Proper selection of the thermal mass, heater size, control electronics and thermal isolation could make this requirement achievable.

2.3 SHADING

As second generation FLIR systems achieve extremely good sensitivity, shading in the displayed imagery will become more noticeable. These shading effects are a result of the apparent background temperature differing from the actual background temperature as a function of FLIR scanner angle. (See Reference 2.2 for a detailed analysis of these effects).

Given a perfect opto-mechanical scanning system, all of the IR energy admitted through the detector entrance aperture originates from the world outside the FLIR sensor. In reality, in any achievable opto-mechanical scanning systems some of the energy which is admitted through the detector entrance aperture originates from the FLIR sensor itself. This energy results from vignetting, reflectivity, and emissivity in the IR telescope. The amount of shading caused depends upon the difference in temperature between the background and the FLIR sensor.

Vignetting can be eliminated by proper optical mechanical design.

Shading caused by reflectivity variations, and emissivity variations, in the lens (induced by scan angle) are typically 1% - 2% of the temperature difference between the FLIR and the scene. The shading will be a smooth variation across the field of view, typically of 0.5 to $1^{o}C$ amplitude. Because it is smooth, its effect will generally not be disturbing.

Narcissus (that is reflections at the various lens surfaces of the cold detector) can be reduced to the order of $0.05^{o}C$ - $0.1^{o}C$ by careful design of the detector system and of the optical system. However, the narcissus effect is a relatively abrupt level change at the center of the field and may therefore, be more noticeable (and disturbing) than the smooth shading.

If the sensititivy of the FLIR is such that the narcissus is disturbing, it may be necessary to attenuate its effect by introducing an opposite effect derived, electronically, as a circularly symmetric function (F) of scan position (x, y):

1) $$r = (x^2 + y^2)^{1/2}$$

2) $$F = F(r)$$

Since the magnitude of the narcissus effect is small, a relatively crude correction will provide sufficient attenuation.

REFERENCES

2.1 "Focal Plane CCD for Thermoelectrically Cooled Detector Array", Interim Report, Contract DAAK70-76-C-0250, to the Night Vision Laboratory by the Honeywell Systems and Research Centers, March 1977.

2.2 Radiometric Accuracy in a FLIR System. Irving R. Abel, Honeywell Radiation Center (To be published in Optical Engineering May-June 1977).

SECTION 3

INTERFRAME AVERAGING/IMAGE REGISTRATION

3.1 INTRODUCTION

In many FLIR applications, the target of interest will remain within the field of view for at least several seconds. Since FLIR typically operates at 30 frames/second there exists considerable potential for interframe averaging to improve signal-to-noise ratio and, correspondingly, thermal sensitivity.

For example, in 2 seconds a FLIR will provide 60 independent video pictures of the target scene. If these pictures could be put into registration, and then added together, the video signal-to-noise ratio (NETD) of the target scene would be improved by almost eight (8) times ($\sqrt{60}$). In many cases of practical interest, the scene detail will remain within the field of view much longer than 2 seconds, providing the potential for even greater improvements in thermal sensitivity through integration.

Examples of applications where such enhanced thermal sensitivity would be particularly useful are:

- pilotage, when the terrain-contrast needed for navigation/orientation is washed out by rain or mist
- viewing at long ranges through an attenuating atmosphere

The potential for sensitivity improvement through integration exists because the required time-constant for "scene-information" is often of the order of seconds, yet imaging systems (such as FLIR, TV) must employ frame times of the order of 1/30 second in order to provide the visual effect (optical illusion) of flicker-free moving imagery. By means of integration, in registration, the benefits of a relatively long integration time can be achieved without sacrificing the dynamic benefits of a 1/30 second frame time.

As will be shown, integration can be effected in real time, with moving images, so that:

(1) all scene detail is displayed at the same video gain, no matter how long (or short) a time over which the detail may have been integrated

(2) there is no appreciable delay in the display of the scene image

When a scene element first appears in the FLIR video it is immediately displayed. On the first frame there is, or course, no reduction of noise through integration. However, in successive frames the noise level associated with that particular scene element progressively decreases through integration. The video level of the detail, itself, remains constant.

In order to usefully integrate moving video imagery it is obviously necessary that successive frames be brought into registration.

There are numerous other ways in which successive frames in registration can be processed to provide useful benefits. (For example, subtraction of registered successive frames of moving video imagery will provide a moving-target-indication of small elements of detail moving differentially within the frame). Accordingly, this section is devoted to image registration, with interframe averaging regarded as one (important) application.

3.2 PRACTICAL IMPLEMENTATION

To integrate (or to perform any operation) on successive video frames requires, in the first place, a frame-storage medium. As a result of recent technology developments, digital memory is now the preferred frame-storage medium.

For example, digital conversion of the input analog video can be effected, at a 10 MHz conversion rate to 8 bit words, by the DATEL ADC-UH8B converter. This module is 5" x 3" x 1.5" and costs on the order of $700.

The digitized video can be stored in digital CCD memory. For example, a complete video frame consisting of 512 x 512 pixels at 8 bits per pixel can be stored in only 32 chips using the new 64 K CCD chips (Fairchild CCD 464). It is projected that 256 K CCD chips will be available in the future, reducing the memory chip count to 8. (For some registration applications full-frame storage may not be necessary, so that even fewer memory chips could be used).

The video-frame digital memory is not only very compact (32 chips using 1977 technology) but is also inexpensive. Figure 3.1 illustrates the rapidly reducing cost of storing a 512x512x10-bit FLIR image in digital memory.

The speed of CCD digital memory is currently in the 2 to 5-MHz range. This is adequate even for the highest video rates because the memory is accessed in sequence. Techniques, such as interleaving and double-word length can be used to increase the effective data rate by several times:

INTERLEAVING:

Store word-1 in chip-1, word-2 in chip-2, word-3 in chip-3, word-4 in chip-4, word-5 in chip-1, word-6 in chip-2, etc. Then, word-rate = 4* chip-rate.

DOUBLE-WORD-LENGTH:

Memory has 16-bit word-length.
Then, 8-bit word-rate = 2* memory word-rate.

Digital memory has the very important advantage of unlimited dynamic range. (Dynamic range is specified by the word-length). This is extremely important for integration, because the dynamic range of the integrated medium must be greater than that of the final image. For example, suppose that 256 frames of unity signal-noise video are to be integrated. The LSB of the ADC will be chosen to correspond, approximately, to the video noise level. After accumulating 256 such samples, the video noise level will be 16 (= $\sqrt{256}$) times

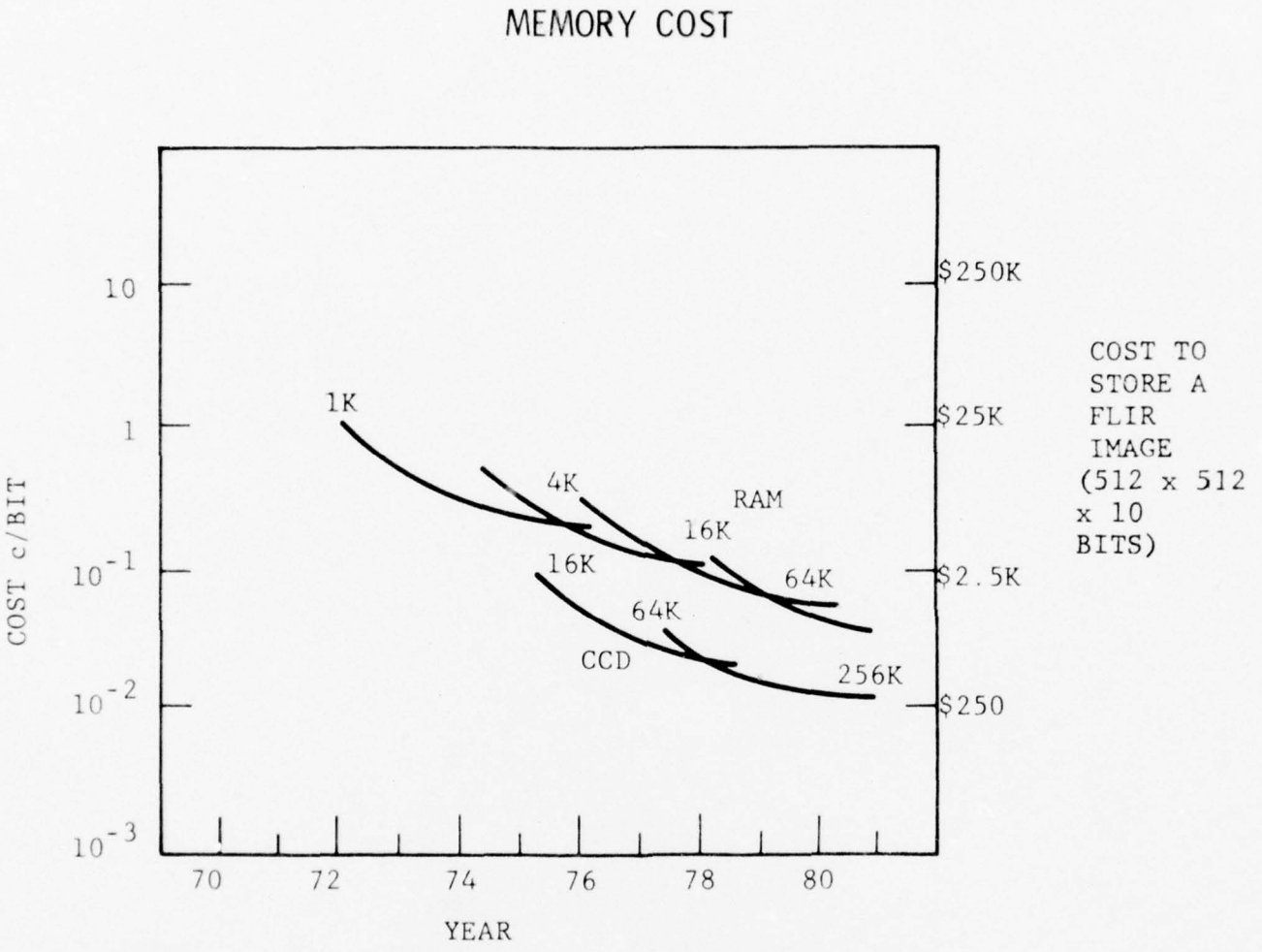


Figure 3.1 COST TO STORE A FLIR IMAGE (ADAPTED FROM ELECTRO '76 PAPER "PROGRESS IN MEMORY AND STORAGE TECHNOLOGY" BY GEORGE FETH OF IBM RESEARCH CENTER)

greater, and the signal 256 times greater, than the level corresponding to the LSB. Although, in this case, the S/N of the output image is only 16:1 ($256/\sqrt{256}$), the integrating-memory is required to have a much greater dynamic range - namely 256:1. In general:

$$(\text{Dynamic Range of Memory}) = (S/N)_{\text{output}} * [(S/N)\ \text{Improvement Ratio}]$$

The technology developments summarized above have demonstrated the hardware feasibility (cost, size) of image storage, and thus of integration, in digital memory. The problem of how to detect and correct in real-time for the frame-frame displacements caused by scene motion (i.e., to effect registration) remains to be considered. At first sight, this would seem to be a particularly difficult problem because of the (apparent) need to perform complex image-correlation operations at very high (video) data rates (of the order of 10 MHz). On closer examination, however, it is seen that the detection and correction of frame-frame displacements can be effected quite easily.

The first factor to be considered is that the scene motion in question is restricted to be that caused by the angular and/or linear velocity of the sensor relative to the scene (Figure 3.2). It is shown in Reference 3.1 that the most general frame-frame displacement (of this kind) can be defined by just eight scalar parameters that are relatively slow-changing. These parameters can be thought of as defining the relative shift, rotation, warping, etc, between successive frames. That is, an 8-parameter distortion of one frame will cause it to be exactly registered with the next frame, by appropriate displacement, rotation, stretching, etc (Figure 3.3).

Secondly, the frame-frame displacement is small and largely predictable. As discussed in Appendix A, the image velocity (for registration applications) is limited so that the maximum frame-frame displacement is $\sim 10\%$ of which the unpredictable component (due to acceleration) is $\sim 1\%$. (The practical significance of these motion limitations is discussed in Appendix A, in relation to typical mission scenarios).

Assuming, for the moment, that the 8 distortion parameters are known, registration of a prior frame (in memory) with the current frame (input video) is effected by Address Modification.

If there was no scene motion, registration would obviously result simply by reading out the old frame from memory in exact sequence (that is, in sync) with the input video. In the presence of scene-motion, the simple memory-address (that would be used in the case of no-motion) need only be slightly modified (according to the value of the 8 motion parameters) in order to effect registration.

This address-modification concept is illustrated in Figure 3.4. The readout from memory is address-modified to bring the stored image in registration with the new (input) image. If the value of the 8 motion parameters (from which the address-modification is derived) is incorrect then the old (readout) image will not be in exact registration with the new (input) image. As illustrated in Figure 3.5, the two images are always compared in order to detect any differences (misregistration) and to appropriately update the value of the 8 distortion parameters.

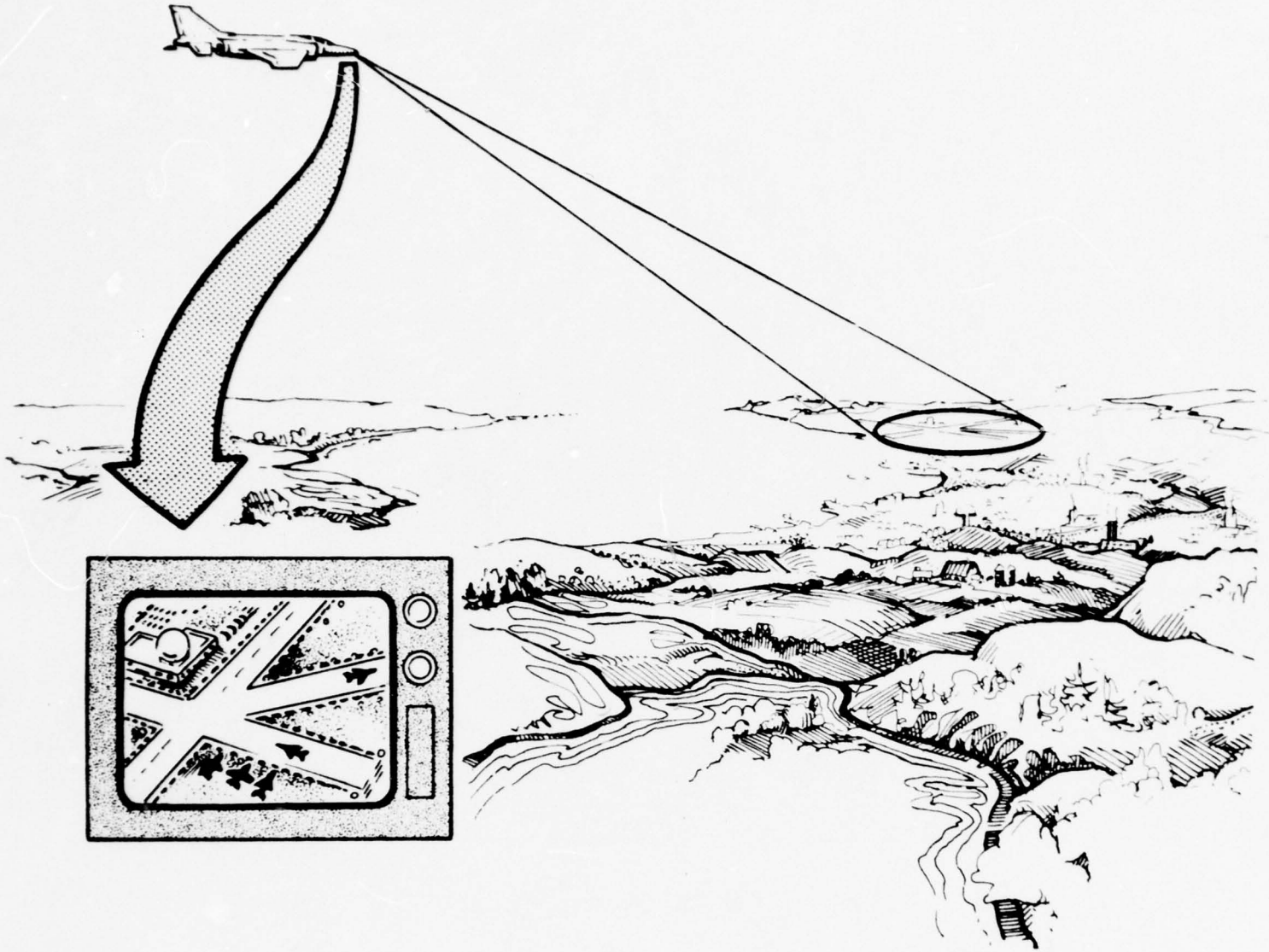


Figure 3.2 MOVING IMAGERY GENERATED BY AIRBORNE VIDEO SENSOR (e.g., FLIR)

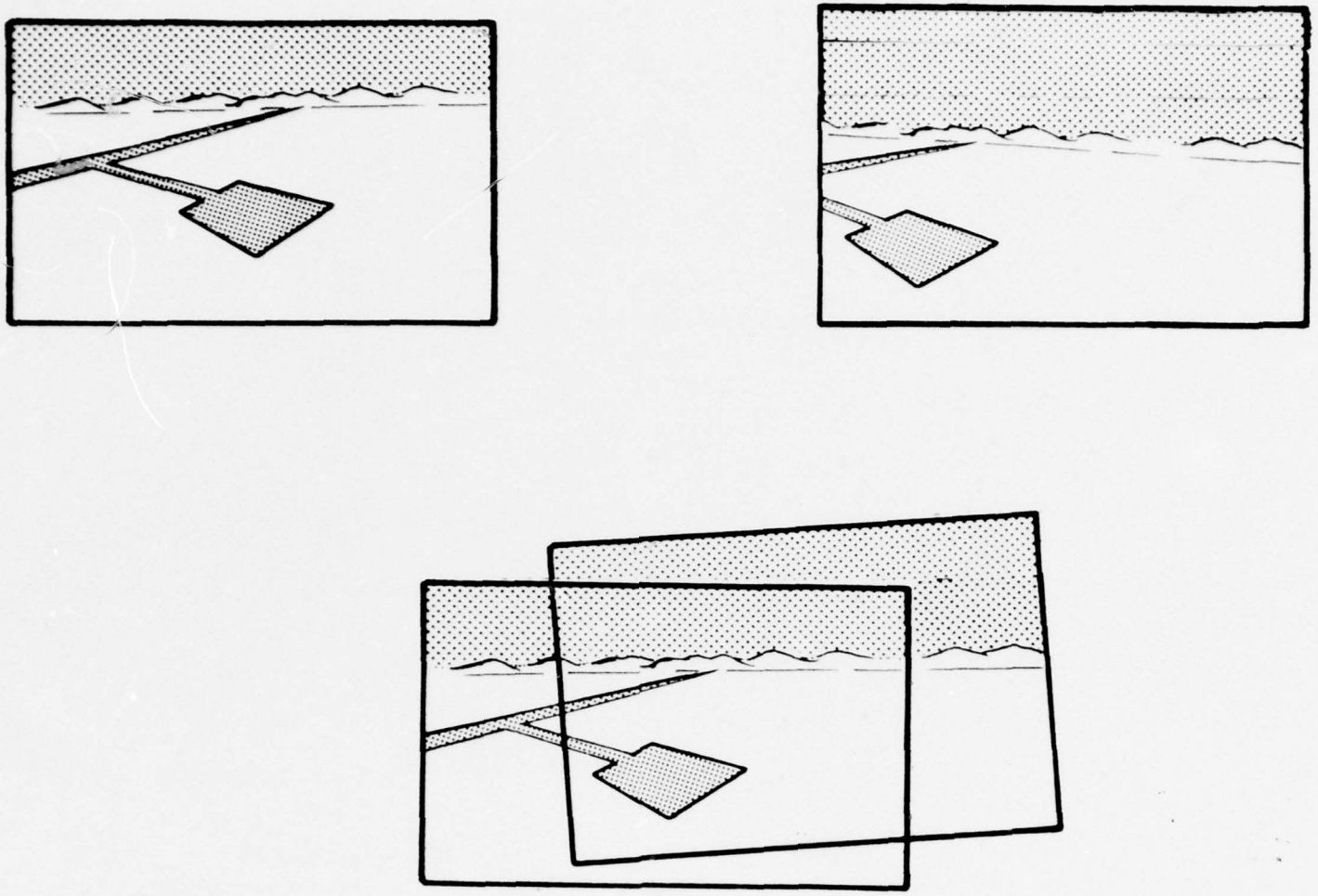
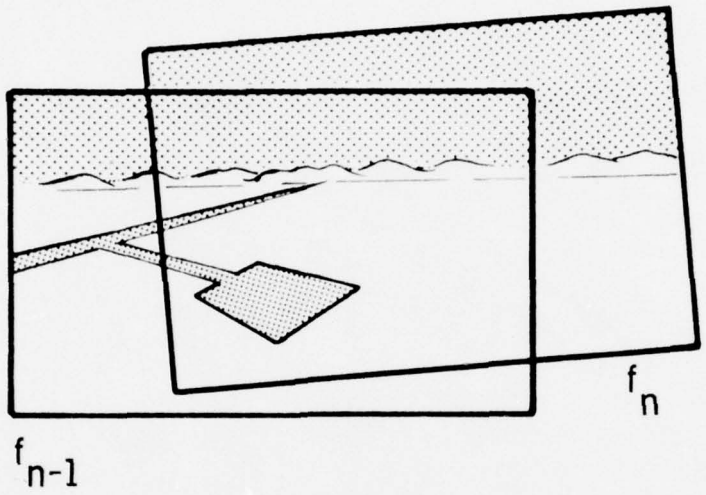


Figure 3.3 TWO SUCCESSIVE VIDEO FRAMES BROUGHT INTO REGISTRATION BY SHIFT/ROTATION/STRETCHING



$$f_n(x,y) = f_{n-1}(x+X, y+Y)$$

WHERE (X,Y) IS ADDRESS MODIFICATION

Figure 3.4 SUCCESSIVE FRAMES BROUGHT INTO REGISTRATION

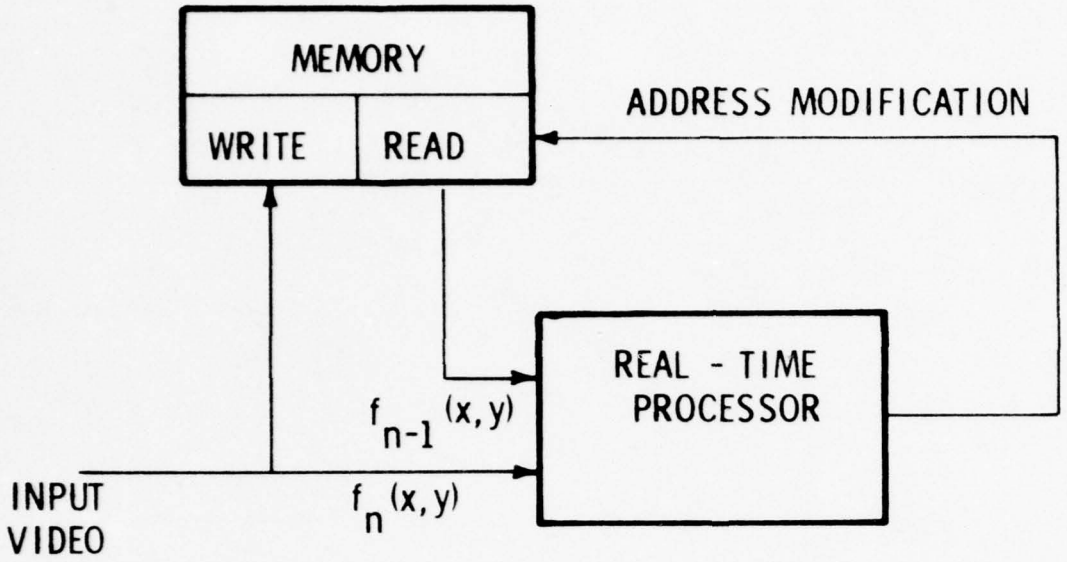


Figure 3.5 MOVING IMAGE PROCESSOR (TO EFFECT REGISTRATION)

The address-modification technique has been described in terms of continuous image functions $f_n(x,y)$, $f_{n-1}(x,y)$. For continuous functions, an address modification (x,y) can be derived which will bring the two frames into exact registration:

$$f_n(x,y) = f_{n-1}(x + X, y + Y)$$

However, in a digital implementation, the fractional pixel component of the address modification cannot be implemented (an image must be registered in one pixel position, or the next, and not at some intermediate interpixel position). Thus, the digital (or discrete) registration is not exact:

$$f_n(i,j) = f_{n-1}(i + I, j + J) + \text{fractional pixel error}$$

The effect of this fractional pixel error is investigated in Appendix B. It is shown that for the integration of successive video frames, fractional pixel errors result in a slight degradation of the image MTF and impulse response. In a practical system example, the effect is shown to be negligible because the random fractional pixel errors, in each of the video frames integrated in memory, tend to average out.

The concept of address modification seems to imply a nonsequential access when reading out the picture elements from the memory. (Sequential access was assumed in the discussion of techniques to increase memory speed such as interleaving and double-word length). There are several approaches to handling this question. For example, the frame-frame distortion is so small ($\sim$ 3% maximum) that the number of out-of-sequence operations will be very small. Thus, even if these particular read operations do take more time (because of the break in sequence) the total extra time involved is very small.

As shown in Reference 3.1, the "warping" of one frame, relative to another is a simple function of scan (x,y) position which can be easily implemented in the memory-read operation. The video comparison operation (to detect and to quantify any residual misregistration) thus becomes the key to the practical solution of the real-time registration problem.

Since the data rate involved is very high ($\sim$ 10 MHz), only the simplest type of operations can be considered. In References 3.1 and 3.2, it is shown how the mean absolute difference (MAD) (between the current frame (f_n) and the previous frame (f_{n-1}):

$$d(xy) = |f_n(xy) - f_{n-1}(xy)|$$

can be processed, in real time, to update the 8 distortion (motion) parameters. This difference function has the obvious property that:

$$D = \iint_{\text{Frame}} d(xy)\, dx\, dy = 0 \quad \text{if frames are identical and are perfectly registered}$$

$$> 0 \quad \text{otherwise}$$

It is this fundamental property which is the basis of the processing operations described in References 3.1 and 3.2.

The squared difference algorithm (SAD) utilizes a similar function:

$$d'(xy) = \left(f_n(xy) - f_{n-1}(xy)\right)^2$$

This has the same fundamental property as the MAD function and can be used without any significant change to the operations that have been referenced. Since high speed ($\sim$ 100 ns) digital-multiplier chips are now available, the SAD function can be implemented just as easily as the MAD function to effect registration.

In the absence of all noise both the MAD and SAD functions will obviously have the same (perfect) registration accuracy.

Tests on a small array (4 x 64) have been performed to assess the correlation accuracy of a number of correlation algorithms when the image has been adulterated with noise.

The results, summarized in Table 3.1, show that the SAD technique is, effectively, as accurate as the best of all the techniques tested. This is important because, being a simple difference technique, it can be easily implemented in real time. (The product-correlation techniques, on the other hand, involve very much more computation and could not be implemented for real-time, full-frame, video processing).

In practice, the difference between the SAD and MAD algorithms may not be significant because the correlation error due to noise would probably be much less, than as indicated in Table 3.1, if the operation is performed over a large (e.g., 512 x 512) array rather than over the small (4 x 64) array used in the comparison tests. In any event, either technique is available for practical real-time implementation as described in References 3.1 and 3.2.

The major hardware items required to implement address modification are illustrated in Figure 3.6.

The current (input frame) is digitized at video rates. The previous frame is read out from memory in synchronism with the input frame. (The Read-Address is developed from the input video sync and the address modification). The current, and the address-modified prior frame, are compared in order to detect any necessary updating of the address modification.

It is an essential characteristic of the processing algorithms described in Reference 3.1 that this comparison operation is pixel-sequential and involves only a relatively few simple arithmetic operations that must be performed at the video rate ($\sim$ 100 ns period). As shown in Figure 3.6, these operations are performed by dedicated hardware. The bulk of the processing operations do not have to be performed at the video rate and can, therefore, be implemented by one (or several) microprocessors.

Table 3.1

EVALUATION OF CANDIDATE CORRELATION ALGORITHMS

1. MAD $J_{p,q} = \sum_{i=1}^{4} \sum_{j=1}^{64} |r_{i+p-1,j+q-1} - s_{i,j}|$

2. SAD $J_{p,q} = \sum_{i=1}^{4} \sum_{j=1}^{64} (r_{i+p-1,j+q-1} - s_{i,j})^2$

3. PROD COR $J_{p,q} = \sum_{i=1}^{4} \sum_{j=1}^{64} r_{i+p-1,j+q-1} s_{i,j}$

4. Canonical Product Correlation (CPC)

$$J_{p,q} = \frac{\frac{1}{n}\sum_{i=1}^{4}\sum_{j=1}^{64} r_{i+p-1,j+q-1} s_{i,j} - \left[\frac{1}{n}\sum_{i=1}^{4}\sum_{j=1}^{64} r_{i+p-1,j+q-1}\right]\left[\frac{1}{n}\sum_{i=1}^{4}\sum_{j=1}^{64} s_{i,j}\right]}{\left[\frac{1}{n}\sum_{i}\sum_{j} r^2_{i+p-1,j+q-1} - \left[\frac{1}{n}\sum_{i}\sum_{j} r_{i+p-1,j+q-1}\right]^2\right]^{\frac{1}{2}}\left[\left[\frac{1}{n}\sum_{i}\sum_{j} s^2_{i,j} - \frac{1}{n}\sum_{i}\sum_{j} s_{i,j}\right]^2\right]^{\frac{1}{2}}}$$

where $n = 4 \times 64 = 256$

5. Squared CPC (square of above)

6. Squared CPC with denominator simplification

$$J_{p,q} = \frac{\left[\frac{1}{n}\sum_{i}\sum_{j} r_{i+p-1,j+q-1} s_{i,j} - \left[\frac{1}{n}\sum_{i}\sum_{j} r_{i+p-1,j+q-1}\right]\left[\frac{1}{n}\sum_{i}\sum_{j} s_{i,j}\right]\right]^2}{\frac{1}{n}\sum_{i}\sum_{j} r^2_{i+p-1,j+q-1} - \left[\frac{1}{n}\sum_{i}\sum_{j} r_{i+p-1,j+q-1}\right]^2}$$

7. Squared CPC with unity sample mean variance

$$J_{p,q} = \frac{\left[\frac{1}{n}\sum_{i}\sum_{j} r_{i+p-1,j+q-1} s_{i,j} - \frac{1}{n}\sum_{i}\sum_{j} r_{i+p-1,j+q-1}\right]^2}{\frac{1}{n}\sum_{i}\sum_{j} r^2_{i+p-1,j+q-1} - \left[\frac{1}{n}\sum_{i}\sum_{j} r_{i+p-1,j+q-1}\right]^2}$$

8. Polarity coincidence correlator (PCC)

$$J_{p,q} = \sum_{4}\sum_{64} \text{sgn } r \text{ sgn } s$$

EFFECT OF NOISE ON CANDIDATE ALGORITHMS

4 x 64 SAMPLE ARRAY

SNR = 2.2 dB

TRIAL CASE	CANDIDATE ALGORITHM ERRORS					
	MAD	SAD	PC	CPC	SCPC	SCPC-DS
Average (over 17 trials) Error	4.40	1.87	6.14	1.73	1.68	1.85
CEP (cells)	4.13	1.76	5.76	1.62	1.58	1.74

The squared CPC with unity sample mean variance, and the Polarity Coincidence Correlator (PCC) were both eliminated during the tests because of poor performance.

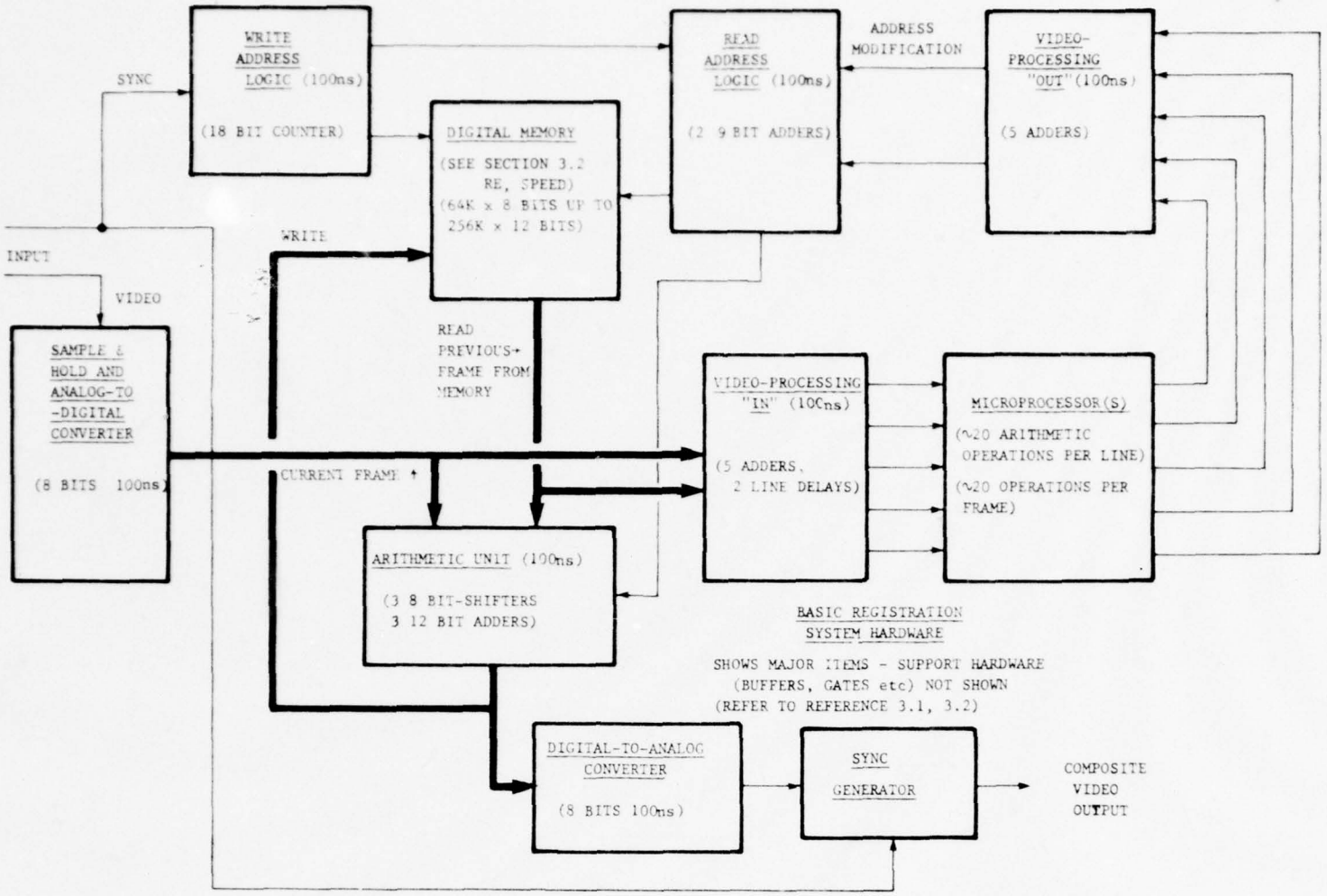


Figure 3.6 HARDWARE IMPLEMENTATION OF INTERFRAME AVERAGING

3.3 REGISTRATION APPLICATIONS

In this section, a number of potential applications of image registration will be described (Figure 3.7).

3.3.1 Enhanced Thermal Sensitivity

This application was introduced in the last section. The potential improvement on thermal sensitivity, as a function of integration time, is shown in Figure 3.8, assuming a 30-frame/s video system.

The thermal sensitivity improvement factor, due to integration, is less for the MRT than for the NETD. This is because some temporal integration ($\sim$0.1 s) takes place in the human eye itself, and the corresponding gain in sensitivity for this is already taken up in the MRT (a subjective measure) but not in the NETD (an objective measure). (0.1 s to 0.25 depending on luminance-0.135 best overall approximation. See Reference 3.5).

When a scene element first moves into the video frame, there is no corresponding picture element in the memory (of the integration system) to which the input video can be added to effect integration. Accordingly, the first time a scene element occurs, it is simply written into memory at some operator specified gain (G):

$$\text{memory} = G * \text{input}$$

As this scene element continues to move through the frame, its input video representation is accumulated with the prior memory samples:

$$\text{(new) memory} = \text{(old) memory} - \frac{\text{(old) memory}}{H} + \text{input}$$

where H is an operator-specified gain factor.

This integration law causes the video level in the memory to rise up to H times the input video level with an (RC) time constant of H frames.

Normally, the operator would choose G = H. Then, as soon as detail moved into the frame it would be immediately displayed at its full (final) video level. The signal-to-noise ratio in this first display of the detail would be equal to the input video signal-to-noise ratio.

On succeeding frames, the scene detail would continue to be shown at the same video level, with the associated noise level decreasing through integration.

No matter for how many, nor for how few, frames a particular scene detail has existed within the field of view, it is always displayed at the same video level. The associated noise level decreases, however, through integration over the time that the detail has existed in the frame or (if shorter) the selected integration time (H fields).

There is negligible delay (1/30 s) between detail first appearing within the field of view and its appearance on the operator's display. There is, therefore, no loss of reaction time through the use of the integrator, only a gain in sensitivity (which develops with time).

- ENHANCED THERMAL SENSITIVITY
- IMPROVED FOCAL PLANE DYNAMIC RANGE
- REDUCED DATA RATE FOR AUTO-CUEING
- REDUCED DATA AND SCAN RATE IN FLIR
- BANDWIDTH REDUCTION (FOR RPV)
- AUTOMATIC GAIN AND OFFSET ADJUSTMENT
- AUTO-FOCUS (SENSE FOCUS MODULATION)
- IMAGE STABILIZATION
- MOVING TARGET INDICATION
- CONTINUOUS DISPLAY/INTERMITTENT VIDEO
- ZOOM
- PYROELECTRIC PROCESSING
- MODULARITY

Figure 3.7 APPLICATIONS OF REGISTRATION

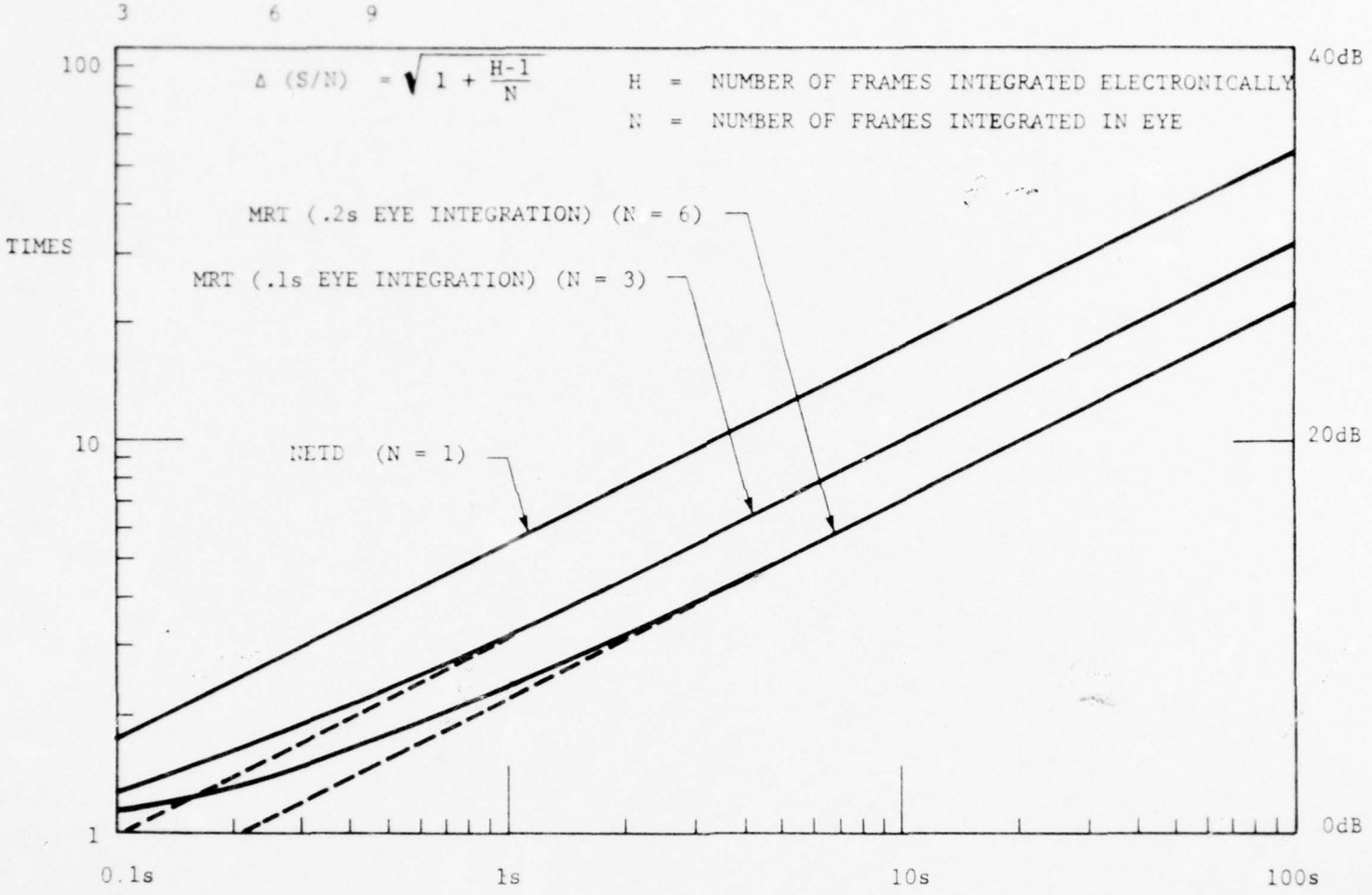


Figure 3.8 GAIN IN THERMAL SENSITIVITY THROUGH ELECTRONIC IMAGE INTEGRATION (30 fps SENSOR)

These benefits of integration apply, of course, only to those scene elements which do not move within the scene during the specified integration time. That is, when the apparent scene motion is due to the relative motion of the sensor only.

(When the target of interest is moving within the scene - for example a truck moving along a road - the registration technique can be applied to enhance the moving target through cancellation of the "stationary" background. See Section 3.3.9).

3.3.1.1 Integration In Sensor

It has been seen how video frames, from a 30-frame/second sensor, can be registered and integrated in digital memory to provide improved sensitivity (Figure 3.9). A similar concept (Figure 3.10) can be used when the integration is performed in the sensor itself by virtue of a lower frame rate (e.g., 1 frame per second). The same type of registration electronics now serves, not to integrate the video, but to generate from each input frame the high-frame-rate (30 fps) sequence of flicker-free moving-imagery required by the operator.

Integration in the sensor is slightly less effective than integration in memory because the noise component of the displayed image is identical in each set of the synthesized frames. (Δ (SNR) = $\sqrt{H}$, integration in sensor; Δ (SNR) = $\sqrt{n + H - 1}$, integration in memory - where the number of frames integrated is H and n is the number of displayed frames per eye integration time. n is typically 3-6 so there is an effective integration loss of between 2 and 5 frames when the integration is done in the sensor).

The potential benefits of this approach are:

- low video bandwidth (for RPV)
- low scan rates (simpler FLIR)
- low video data rate (for auto-cueing)

These applications are considered in more detail in subsequent sections.

Six potentially deleterious effects of reducing the frame rate of an image sensor can be identified:

(1) flicker (of all detail)
(2) smear
(3) loss of picture area
(4) distortion, or loss of, picture detail
(5) flicker of differentially moving objects
(6) degraded tracking accuracy of an optical tracker.

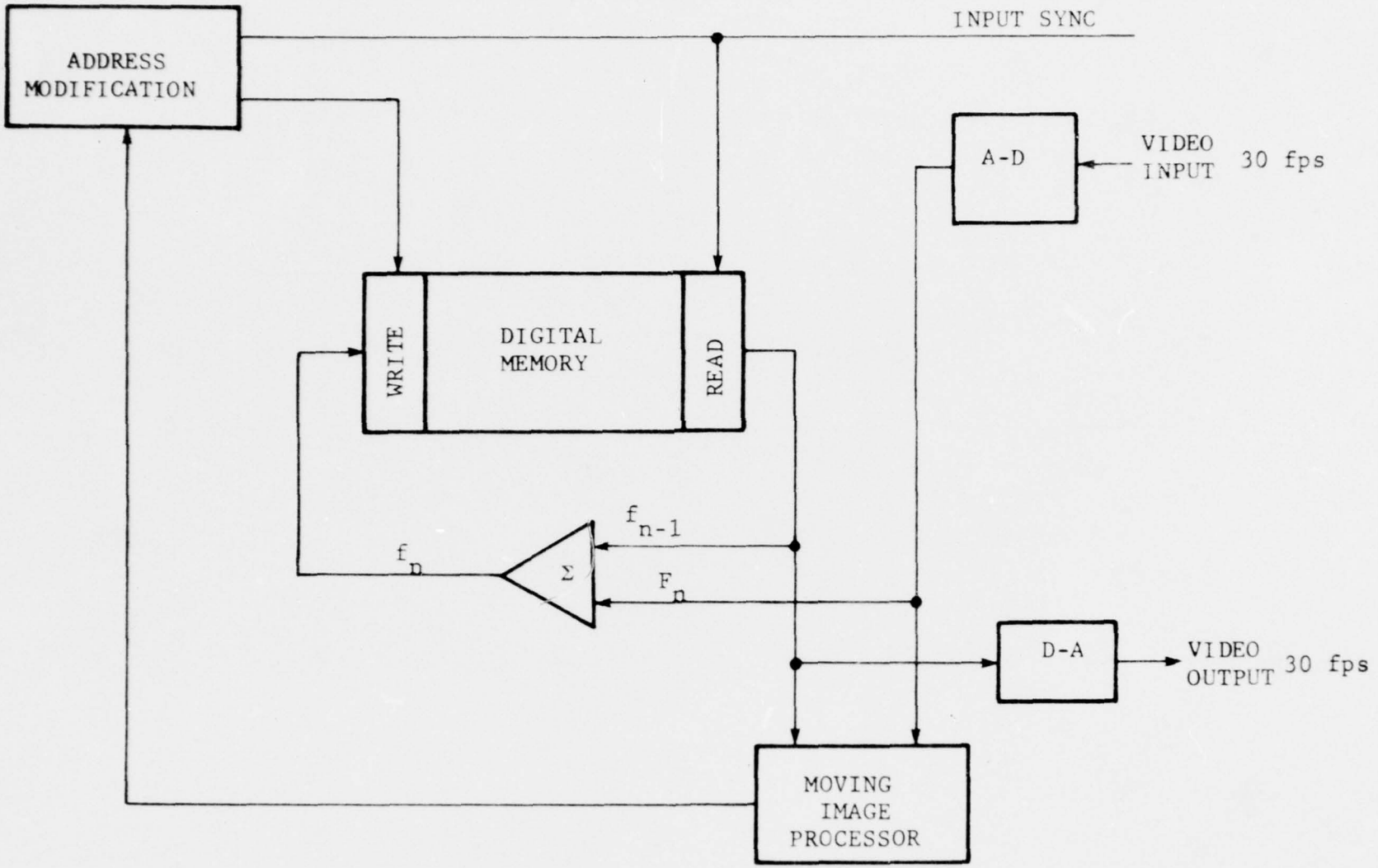


Figure 3.9 DYNAMIC IMAGE INTEGRATION/INTEGRATION IN MEMORY

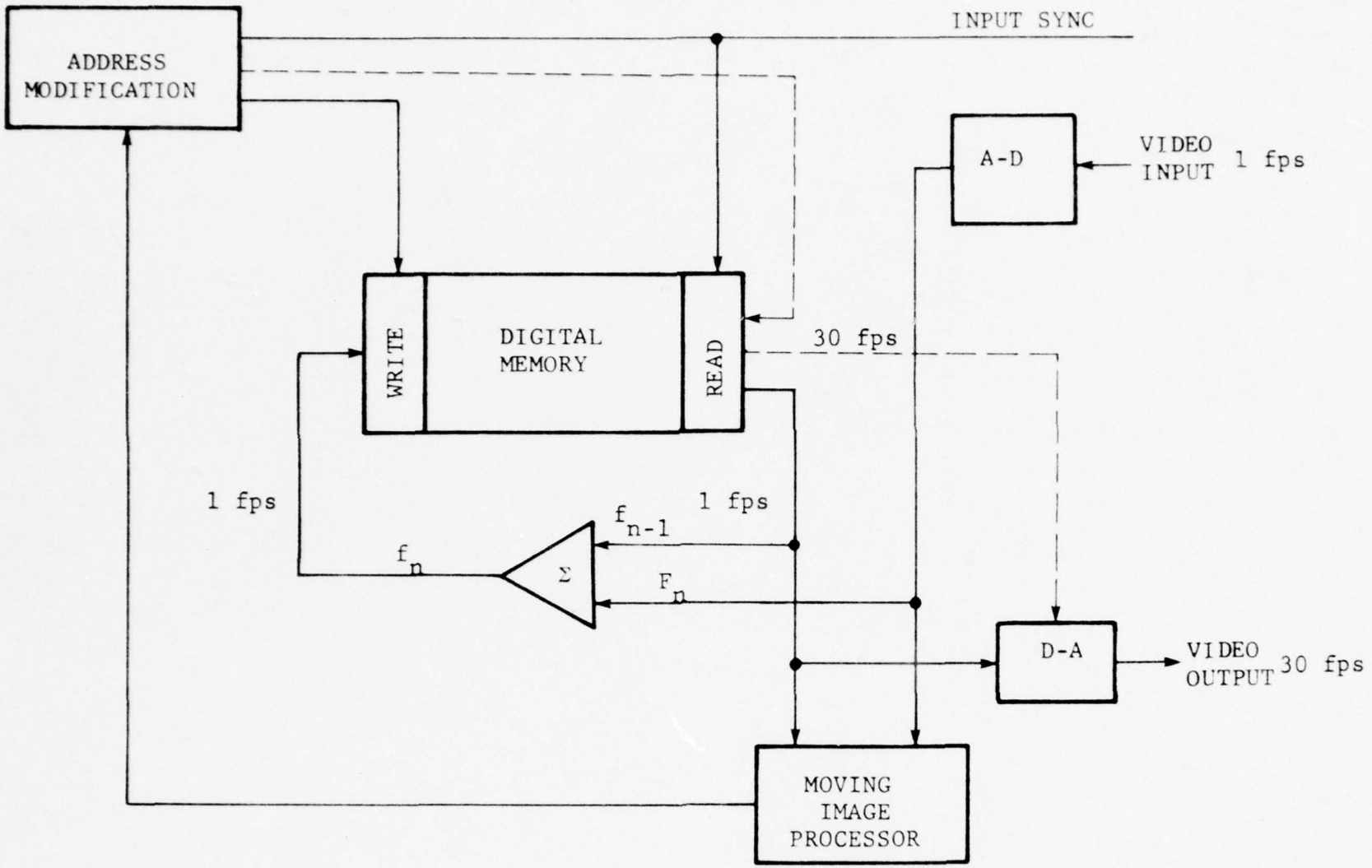


Figure 3.10 DYNAMIC IMAGE INTEGRATION/ INTEGRATION IN SENSOR

As noted above, the registration electronics can synthesize 30 properly displaced ("moving") frames per second for display from a smaller number of sensor-frames per second. In this way the flicker effect of the reduced frame rate, is avoided.

In the case of a staring sensor, there is a loss of resolution (due to smear) when the image moves by more than one element in a frame time. Thus at 30 fps, image velocity must be less than 1 element in a 1/30 s to avoid smear. For a 750 x 1000 picture (as illustrated in Figure 2.1) this translates to an image velocity of 3% of screen width/s. At 1 fps the smear threshold is an image velocity of 0.1% of screen width/s. This smear threshold as a function of image velocity is illustrated in Figure 3.11. It shows that a high frame rate is needed - not just to avoid flicker - but also to avoid smear. Thus the flicker-free reduced frame-rate concept is not (generally) applicable to staring sensors.

If the image velocity is so high that there is little or no overlap between sensor frames then clearly the frame synthesizer will be unable to show a continuous full-frame view (see Figure 3.12). It can be seen that to avoid this effect the image motion, per sensor frame, should be limited to the order by 10% of screen width. This ("Loss os Picture Area") limitation is also shown in Figure 3.11.

In addition to loss of picture area (in synthesized frames between sensor frames), image velocity can cause loss of detail and/or spatial distortion depending upon the particular type of scanner utilized (serial, parallel, serial-parallel, etc). Again, the image velocity should be limited to 10% of screen width per frame in order to avoid this fourth effect of reduced frame rate.

Figure 3.11 shows that reduced frame rate may be feasible for scanning imagers, depending upon the magnitude of scene velocity to be encountered. (Over to range 10% to 300% screen width/s). It should be noted that scene image velocities will, in practice, be limited by the ability of the human operator to utilize such rapidly moving imagery (even when it is "perfectly" displayed).

Figure 3.11 shows that the movement limitation of a scanning imager at 3 fps is much less than the movement limitation of a staring sensor at 30 fps. This indicates that the image motion limitations (with a 3 fps sensor) may not be of practical significance in many applications.

The previous effects all arise with overall scene motion. If there is differential motion within the scene (for example a truck moving along a road) than there is an additional effect. The differentially moving target will appear to move discontinuously (it will flicker- or jump- at the sensor frame rate). For example, suppose that a 12-foot truck moving at 30 mph is being viewed by a sensor having a field of view (at the truck) of 1200 feet. At 30 sensor-frames per second, the truck moves 44/30 ft, or 12% of its length, per sensor frame (1/30 s). At 1 sensor-frame per second, the truck will appear to jump 44 ft, or 4 times its length, every second.

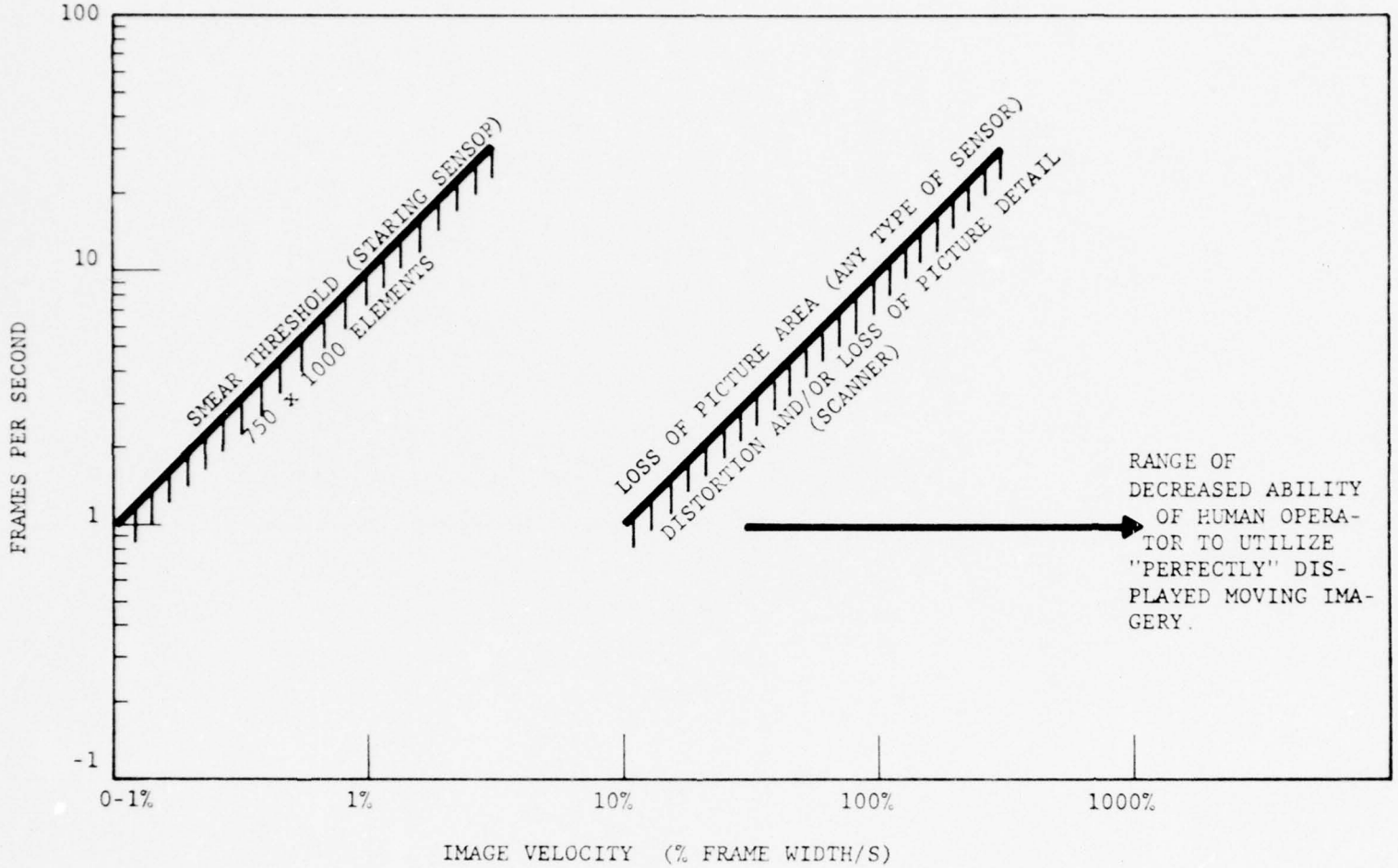
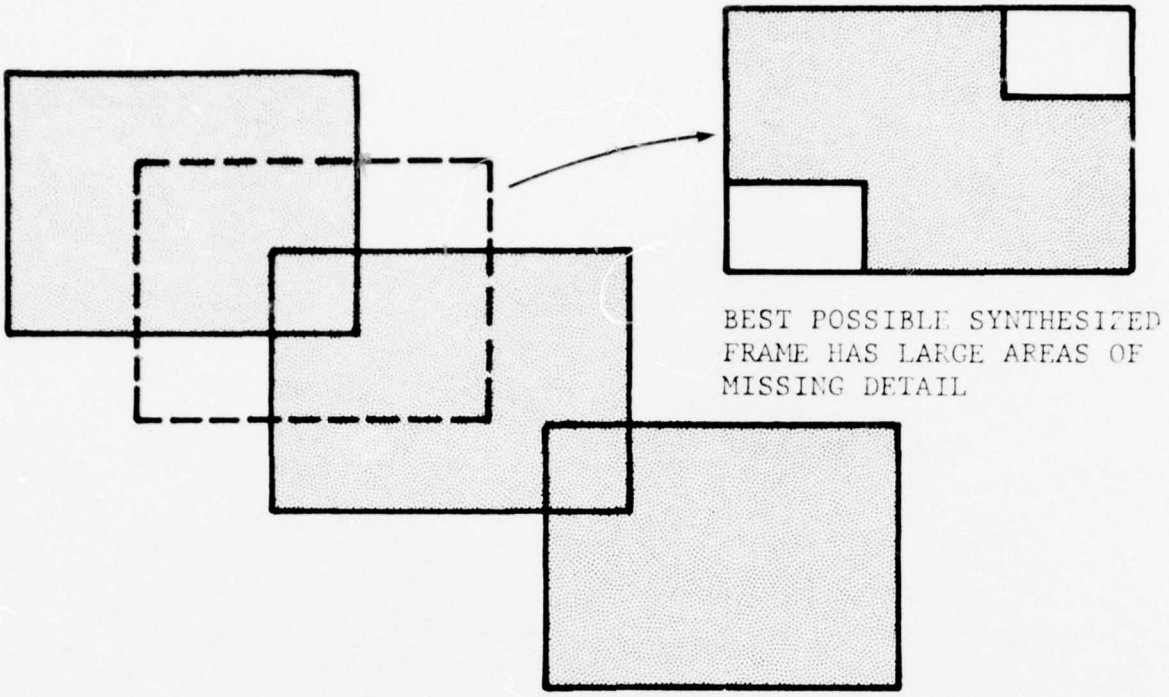
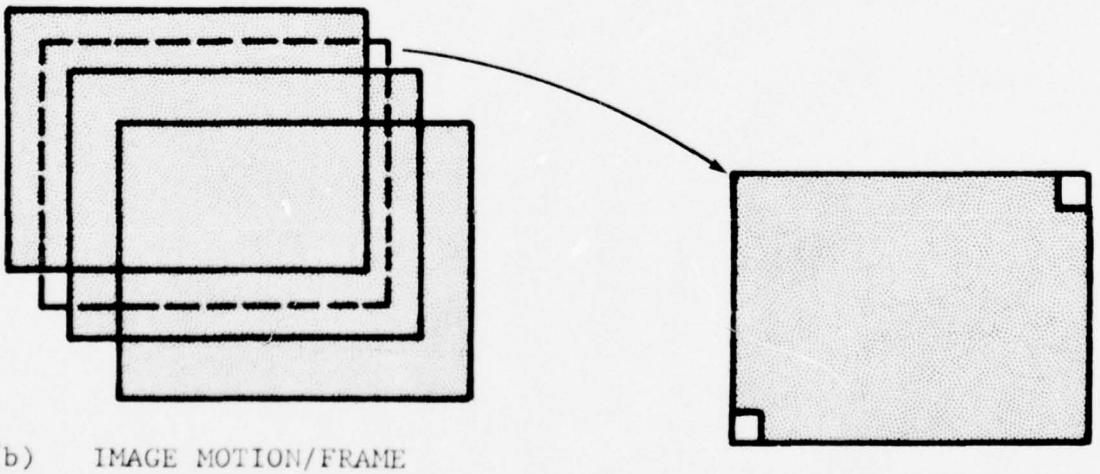


Figure 3.11 LIMITATIONS ON IMAGE VELOCITY



(a) IMAGE MOTION/FRAME ~ 1 FRAME WIDTH



(b) IMAGE MOTION/FRAME
≤ 10% OF FRAME WIDTH

Figure 3.12 LOSS OF PICTURE AREA IN SYNTHESIZED FRAMES

In some applications this moving target effect would be a distracting limitation. In other applications, however, it would provide a useful MTI feature.

Finally, a reduction in the sensor frame rate (e.g., 30 fps to 3 fps) would proportionally decrease the tracker-loop-bandwidth of any associated optical tracker. This would decrease the tracking accuracy with rapidly moving targets. Thus reduced frame-rate would not generally be applicable to image sensors servicing a high speed optical tracker.

3.3.2 Improved Focal Plane Dynamic Range

The required well-capacity of a CCD focal plane readout device is proportional to the square of the specified NETD. In some cases, the desired NETD may require a CCD well-capacity that is greater than can be provided in the CCD. However, if image integration is utilized to help achieve the desired system-NETD (or MRT), then the NETD per frame is reduced, and consequently the required CCD well-capacity is also reduced.

In effect, the scene photons can be considered as being integrated off-the-focal plane (in digital memory) rather than in the CCD charge wells.

The reduction in required CCD well capacity, if image integration is used to meet a given NETD (or MRT) specification, is shown in Figure 3.13 as a function of the integration time utilized.

3.3.3 Reduced Data Rate for Auto-Cueing

Operational requirements for the detection of small targets at ever increasing ranges, are forcing FLIR systems design towards larger numbers of pixels per frame.

This complicates the FLIR design through increased data rates. The biggest problem, however, is at the man/machine (display) interface:

1. As the number of pixels in the picture increases, it becomes increasingly difficult to provide a display on which the man can even resolve the individual pixels. This is principally a matter of the physical size of the display, which is generally severely limited by practical constraints such as cockpit size and layout.

2. Even if the display is made large enough for the man to resolve individual pixels, his mental ability to utilize all of the picture data within a given (short) time falls off as the quantity of the picture data (pixels/frame) increases.

These considerations lead to the conclusion that, as the number of pixels/frame is increased (to meet the increasingly stringent operational requirements), the need for some form of automatic target cueing increases correspondingly:

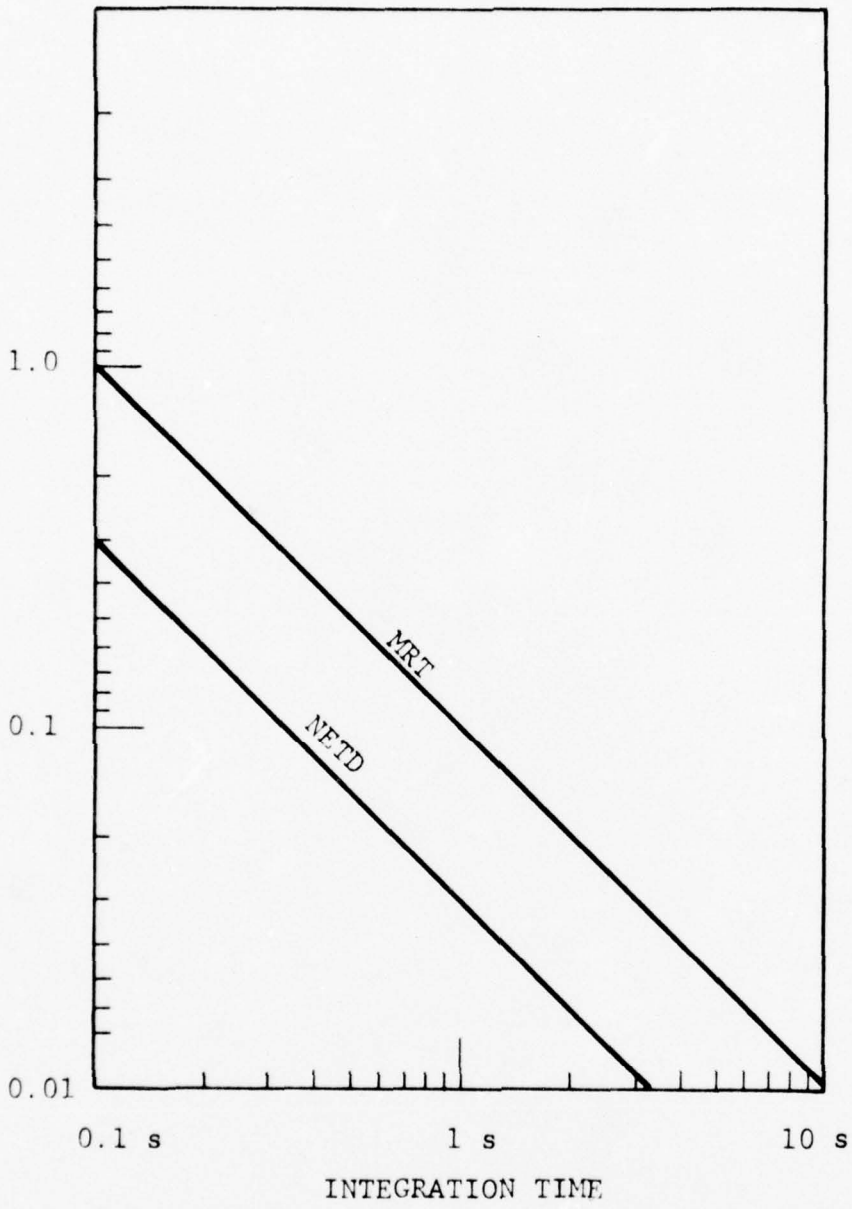


Figure 3.13 RELATIVE CCD WELL CAPACITY REQUIRED IF INTEGRATION USED TO ACHIEVE DESIRED THERMAL SENSITIVITY

With auto-cueing it may not be necessary that the operator be able to resolve all of the pixels in the image generated by the sensor. In a target acquisition (search) mode the display can show the full video frame at normal TV quality and not at the full pixel density inherent in the sensor video.

The full pixel-density sensor video is applied to the automatic target detection (auto-cueing) system. Whenever a potential target is detected, the operator is automatically cued, and may (for example) then examine the indicated (small) target area at full pixel-density using his standard TV quality display by electronically switching, or zooming, his display field of view.

An automatic target cueing system can thus overcome the problems, noted earlier, that arise when operational requirements dictate very high numbers of pixels/frame. Unfortunately, as the number of pixels/frame increases, the video data rate also increases -- making it increasingly difficult to implement automatic target detection.

For example, consider, as a baseline system, a conventional (first-generation) 500 x 500 pixel, 30 frame/second, FLIR sensor. The data (pixel) rate is approximately 7.5 MHz (Figure 3.14).

In a second-generation system, suppose that the required detection range (and field width) has been increased by three times (Figure 3.15). The pixel density is thereby increased to 1500 x 1500 per frame, leading to the necessity for autocueing for fast target search and detection. However, at 30 fps the data rate of the 1500 x 1500 pixel sensor is 67.5 MHz. Auto-target detection is much more difficult at this data rate than at the 7.5 MHz rate of the (first-generation) 500 x 500 pixel sensor. The design challenge presented, therefore, is that the more auto-cueing is needed the more difficult it becomes to implement.

The potential application of the registration electronics is to reduce the video data rate by eliminating the (present) absolute requirement that the sensor generates at least 30 frames/s.

Suppose that a totally automatic target detection system were to be designed utilizing a FLIR-type of sensor and autocueing electronics, but with no operator display of the FLIR image. In this case the frame rate of the sensor would logically be determined through tradeoff between:

- desired reaction time (that is, the time to detect a just appeared target)
- velocity of scene motion
- cost/complexity/feasibility of the auto-cueing electronics as a function of the data rate.

It is unlikely, given that the subjective requirement for a 30 fps sensor is eliminated, that these objective requirements would lead to a 30 fps specification for the sensor frame rate.

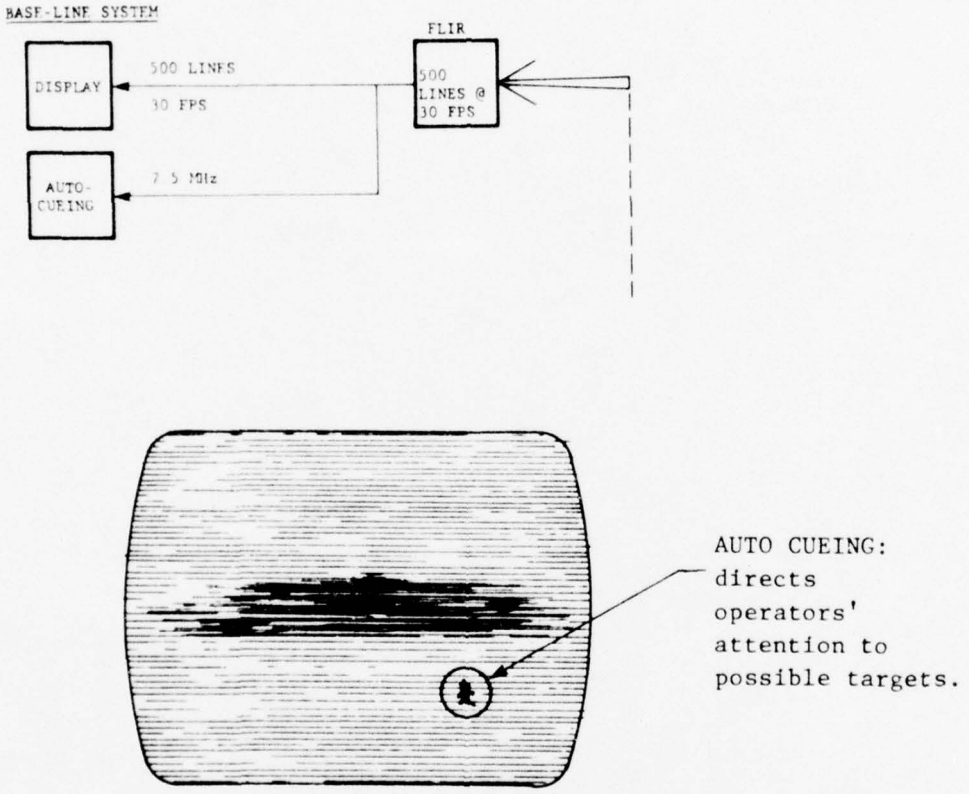


Figure 3.14 AUTO-CUEING SYSTEM FOR FIRST GENERATION FLIR

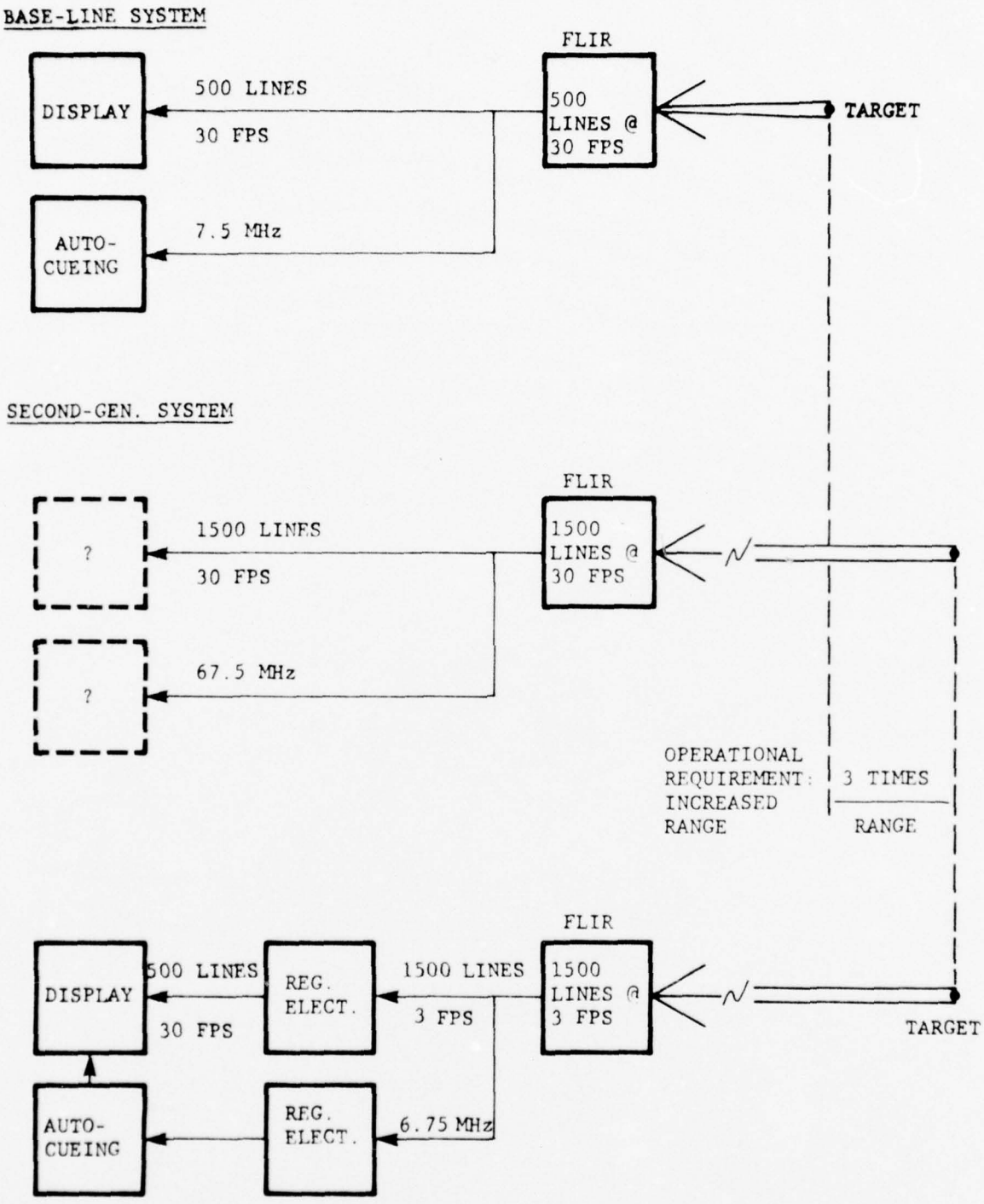


Figure 3.15 DATA-RATE-REDUCTION FOR SECOND GENERATION FLIR AUTO-CUEING

For example, if:

required reaction time ~ 0.3 s

and,

scene-motion velocity $\leq$ 30% of screen width/s

then a frame rate of 3 per second might be adequate.

The video data rate of a 1500 x 1500 pixel second generation sensor could be reduced in this way from 66.5 MHz (at 30 fps) to 6.75 MHz (or 3 fps), less (even) than in the first generation baseline system.

The reduced data rate to the auto-cueing electronics not only simplifies the design of that electronics but also provides increased SNR (through integration in the sensor).

3.3.4 Reduced Data and Scan Rate in FLIR

As previously discussed, operational requirements are leading to increased numbers of pixels/frame - which in turn increases the video data rate. The high video data rate complicates the design of both focal plane, and off-the-focal plane, electronics. However, by reducing the frame rate, the video data rate can be substantially reduced (Figure 3.16)

The particular frame rate of 30 per second arose (historically) mainly to satisfy the flicker-free and scene motion requirements of the viewer. The actual scene-data rate that the operator needs may be much less, for example, of the order of 3 frames/second.

By using the three frames/second sensor, with registration electronics to generate a sequence of 30 properly displaced/stretched/rotated frames per second, the operator's total display-needs may be satisfied with a much lower data rate system.

This simplifies the FLIR electronics. Many more components are available to the designer of a (1000 x 1000 pixel) second-generation FLIR at a 300 ns pixel time (3 fps), than at a 30 ns pixel time (30 fps)

3.3.5 Bandwidth Reduction for RPV

A FLIR sensor may be located on an RPV, and the video signal transmitted back to a base station for display to the RPV operator.

FRAME RATE (FPS) (SENSOR)	30	3
PIXEL TIME	30 ns	300 ns
MIRROR SCAN RATE	30 cps	3 cps
"FLYBACK" TIME (90% EFF)	3.3 ms	33 ms
FRAME RATE (FPS) (DISPLAY)	30	30

Figure 3.16 TYPICAL FLIR DATA, SCAN, RATES FOR A 1000 x 1000 PIXEL IMAGE

It is obviously very important that the video link from the RPV to the base station be as immune as possible to enemy jamming. For this reason, and also to minimize the size and power of the airborne video transmitter it is necessary to use the narrowest possible video bandwidth. There are two ways in which the registration electronics can be applied to achieve bandwidth reduction:

1. The video data is transmitted at a reduced frame rate (e.g., 1 frame/s) and regenerated by the registration electronics (at the ground station) to the full 30 frames per second as required for a flicker-free display of moving imagery

2. A relatively small number (e.g., 70 x 70 $\approx$ 5000) of pixels per frame are transmitted at an intermediate frame rate (e.g., 10 per second). These "mini-frames" are automatically deployed over the scene, according to the interest of the viewer, and are displayed to the operator in the form of the best available collage, in registration, of the current and all prior frames.

The first technique is applicable to standard (e.g., 525) line video and offers bandwidth savings of the order of 30 times (transmitting at 1 frame/s).

The second technique could provide useful scene information with a very much lower video bandwidth (e.g., 20 kHz).

3.3.5.1 Bandwidth Reduction by Frame-Rate Reduction

This technique is potentially applicable when the scene velocity is of the order of 1 frame width per second, or less, so that there is then significant interframe redundancy at 30 frames/second.

The reduced frame rate can be achieved either directly, from a slow-scan FLIR, or with a regular 30 frames/second FLIR, by slow speed readout from an airborne integrator memory (Figure 3.17). (In either case there is an added benefit of improved sensitivity due to integration).

The registration electronics at the ground station performs the function of regenerating 30 properly displaced/stretched/rotated frames per second as required for flicker-free display of moving imagery. (Elements of scene detail that are moving differentially with respect to the scene (e.g., a moving vehicle on a road) will appear to flicker in the regenerated video, thereby, providing a useful MTI feature).

3.3.5.2 Bandwidth Reduction by Selective Transmission

In this technique, each transmitted video frame ("mini-frame") contains a much smaller (than usual) number of pixels -- for example, 5000 pixels ($\sim$ 70 x 70). The frame rate is chosen to be less than 30 per second but significantly higher (e.g., 10 frames/s) than the typical maximum fixation frequency of the viewer (2-5 fixations/s). As will be described further on, the small mini-frames are automatically deployed over various parts of the total scene area.

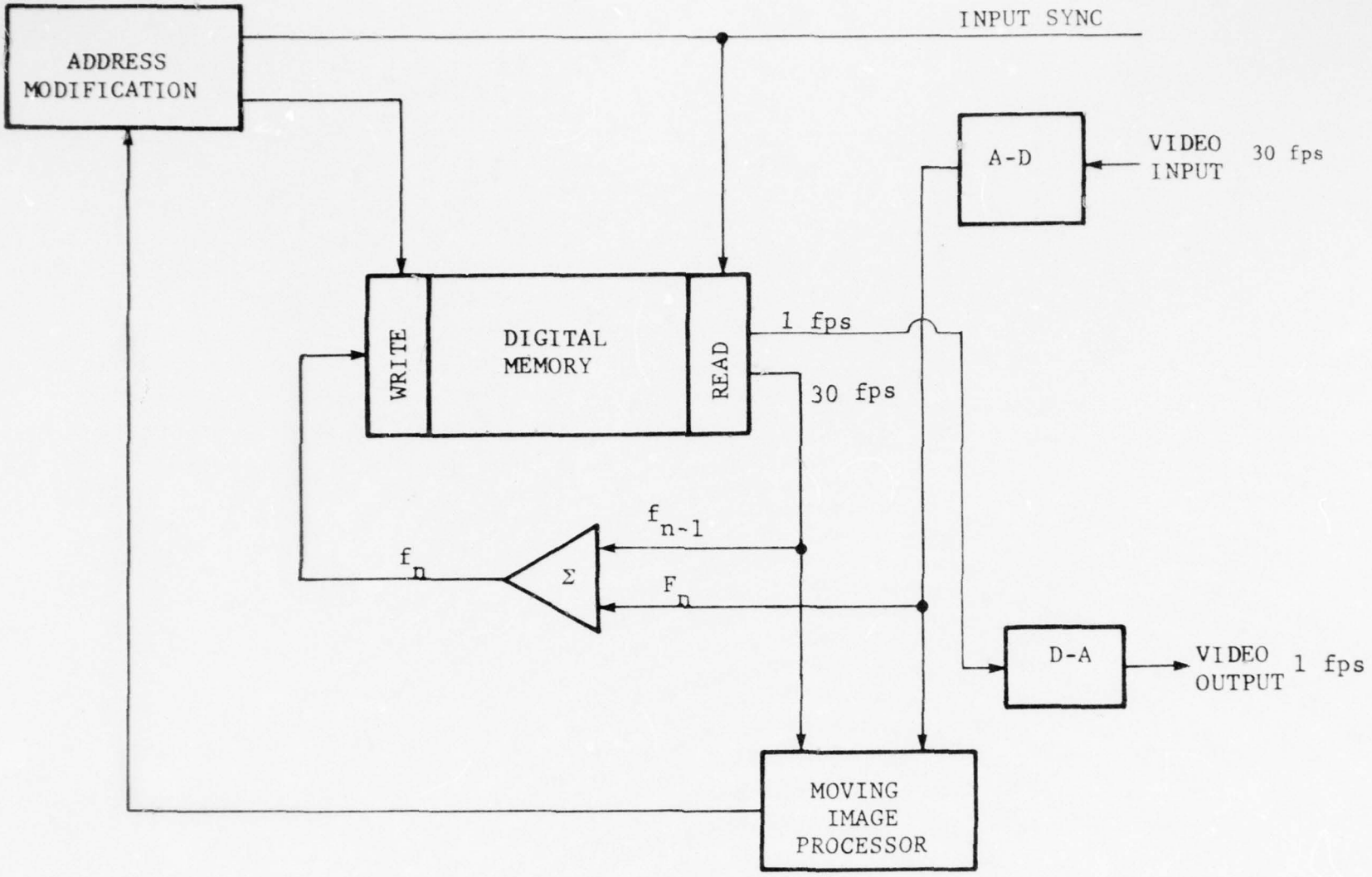


Figure 3.17 FRAME RATE REDUCTION BY SLOW READ-OUT FROM MEMORY

All of the received video frames, showing these various scene sections, are put into registration so as to generate a collage of as much of the scene area as may have been covered in the past by the individual mini-frames.

The operator sees on his display a collage (in registration) of the present, and prior frames, covering the entire target scene area. As the operator looks over this scene, his eye direction is measured with an Oculometer (Figure 3.18), thereby defining the screen coordinates of the area that is of current interest to the operator. This information is transmitted back to the RPV, and is used to cause the next mini-frame to be deployed over the corresponding part of the scene. In this way, the mini-frame is always deployed on that part of the scene that is of current interest to the operator, and where he is currently looking with his foveal vision. The operator's peripheral vision is served by the collage of all of the other (prior) frames.

Initially, the mini-frame is made to cover the entire target scene, at low resolution. This permits the operator to quickly orient himself with the entire general scene area. (This expansion of the mini-frame field of view is done either by zooming-out, or by virtue of the initial (long) range of the RPV from the target scene.)

The field of view of the mini-frame is then made to progressively close in, in order to provide higher spatial resolution over the specific areas that the Oculometer is reporting to be of current interest to the operator. (This compression of the mini-frame field of view is done either by zooming-in, or by virtue of the RPV flying towards the target.)

The potential advantages of this mini-frame technique is that it provides useful imagery (especially for terminal guidance) at an ultra-low video bandwidth (20 kHz). The concept is illustrated in the sequence of Figures 3.19, 3.20, 3.21, and 3.22.

The ground station hardware involved consists of:

- registration electronics (see Figure 3.6) modified as described below
- display monitor with Oculometer
- data link to receive video ($\sim$ 50,000 pixels/s) and to transmit point-of-interest commands (e.g., two, 100 Hz, channels)

The registration electronics, for this application, will utilize address modification of the write (instead of the read) operation. It will also be necessary to scan-convert the low data rate input ($\sim$ 5000 pixels/frame) into a higher density picture (e.g., 250,000 pixels/frame). This could probably best be done by appropriate modification of the digital registration system. Alternatively, it could readily be performed by a storage tube scan converter (such as the PEP-400-R, Princeton Electronic Products, or the TH7502, Thomson CSF).

The digital memory would be of the same size as was shown in Figure 3.6.

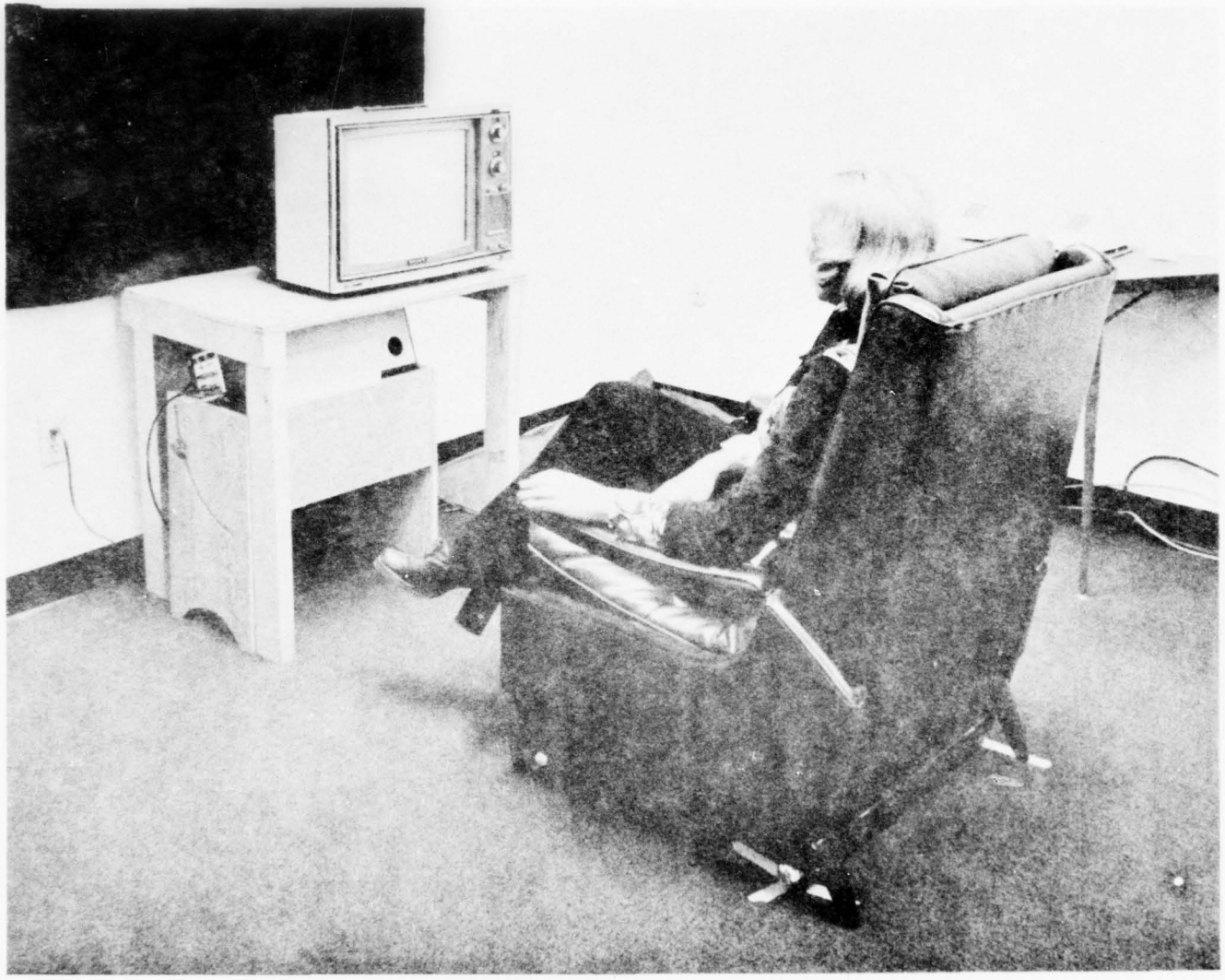


Figure 3.18 HONEYWELL REMOTE OCULOMETER RECORDING FIXATIONS OF SUBJECTS WATCHING TV COMMERCIALS (THE OCULOMETER OPTICAL SENSOR IS SHOWN BENEATH THE TV SET. THE OCULOMETER SIGNAL PROCESSOR (NOT SHOWN) IS A STANDARD MINICOMPUTER 19" x 11" x 24".

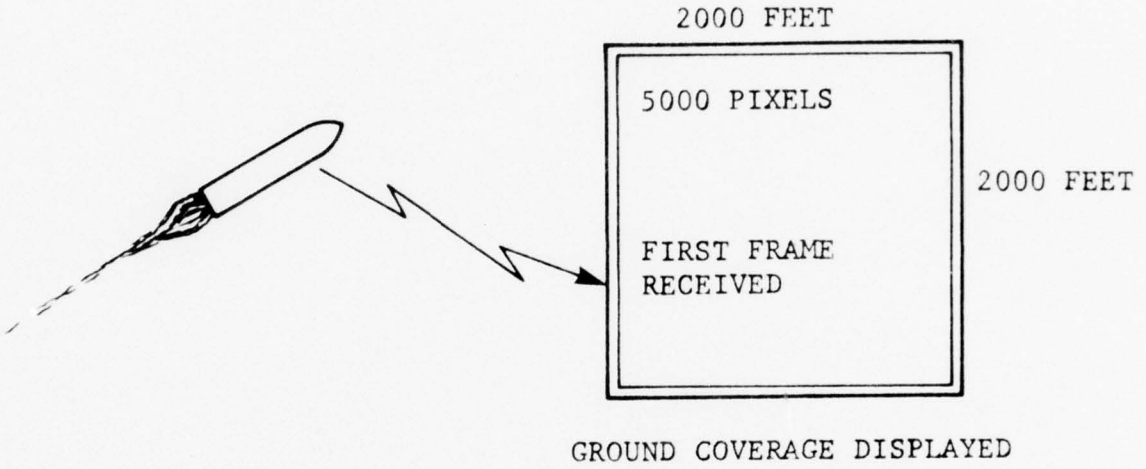


Figure 3.19 30 SECONDS TO GO - MINI-FRAME COVERS ENTIRE TARGET SCENE

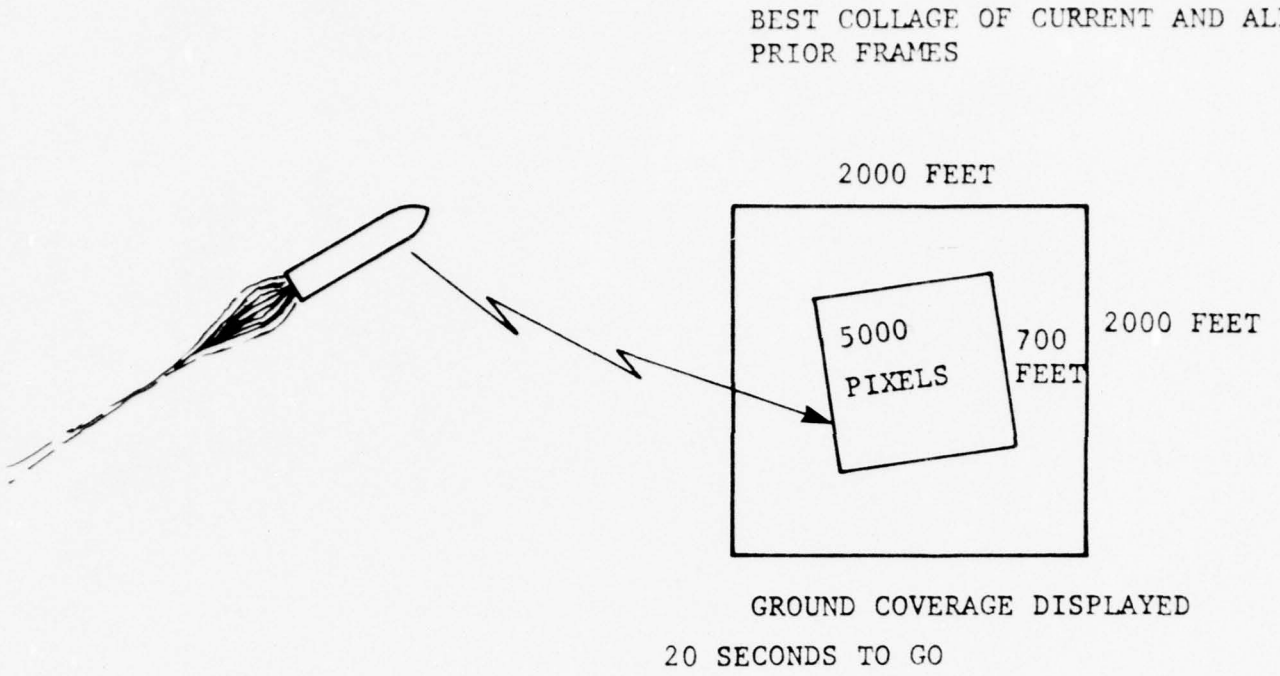
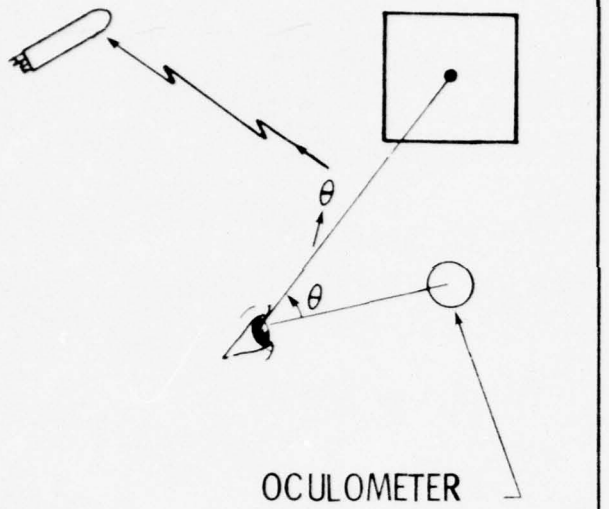
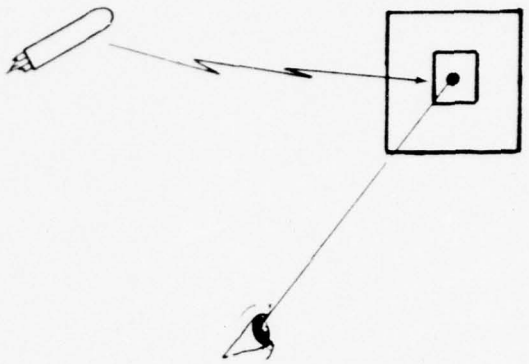


Figure 3.20 MINI-FRAME SHRINKS DOWN TO PROVIDE HIGHER RESOLUTION OF THAT PART OF THE TARGET-SCENE THAT IS OF CURRENT INTEREST TO THE OPERATOR



OCULOMETER
SENSES θ



CURRENT MINI-FRAME IS
(AUTOMATICALLY) ALWAYS
AREA OF INTEREST

Figure 3.21 SELECTION OF CURRENT MINI-FRAME

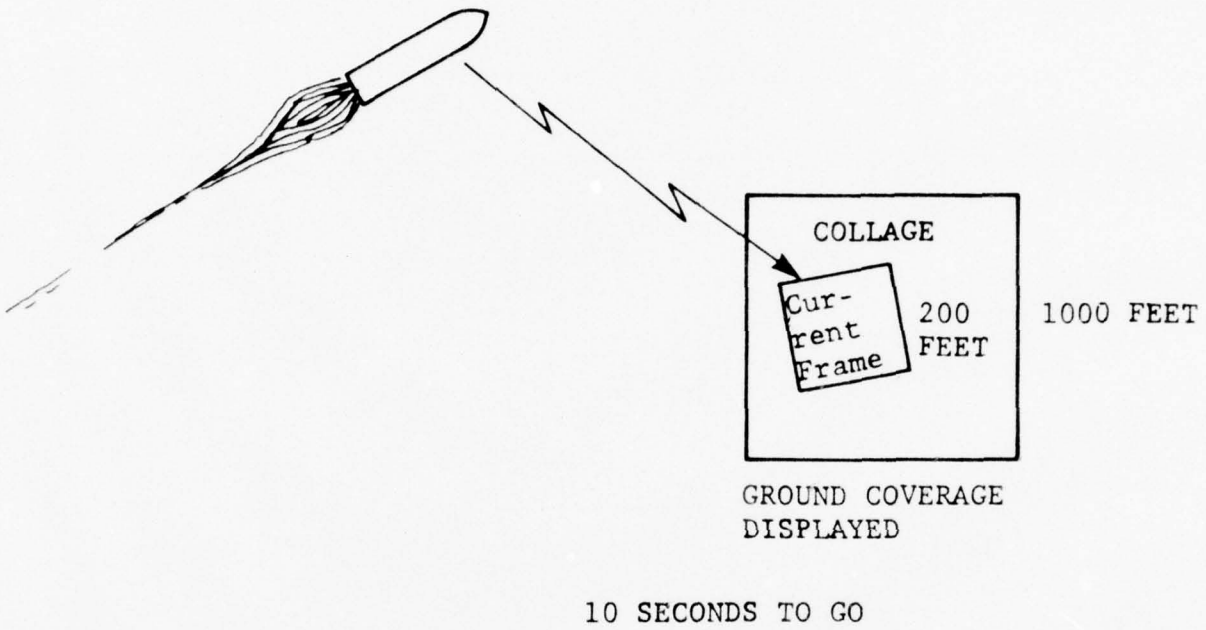


Figure 3.22 COLLAGE AND MINI-FRAME BOTH CLOSE-IN

Thus the registration electronics required for this application may be estimated as to be between 1-1/2 and 2 times that of the conventional registration electronics as shown in Figure 3.6.

3.3.6 Automatic Gain and Offset Adjustment

In Section 2 the scan system of a typical second generation FLIR was described. In this, the scene image is scanned over a linear array of 750 detector elements. A single composite video is generated by multiplexing individual detector outputs.

It was shown how the individual gain and offset variations of the 750 detector channels could be corrected by means of two sets of 750 digital words, circulating in a shift register.

Various calibration techniques were described for generating these 1500 correction words. In this section a possible means is described of generating (or at least fine-tuning) these words automatically, from the video data alone, using the registration electronics.

The general principle of operation of the registration electronics is that a prior frame, stored in memory, can be so distorted (address-modified) that, when read out of memory in synchronism with the input video, it is in exact registration with the input video frame.

It follows that, if there is any vertical scene motion, corresponding lines of these two registered videos must have been generated by two different detectors.

Thus by suitably comparing (over one line) the input video and the stored video, gain and offset corrections can be automatically developed (for the corresponding detector channel) so as to bring the two videos into equality.

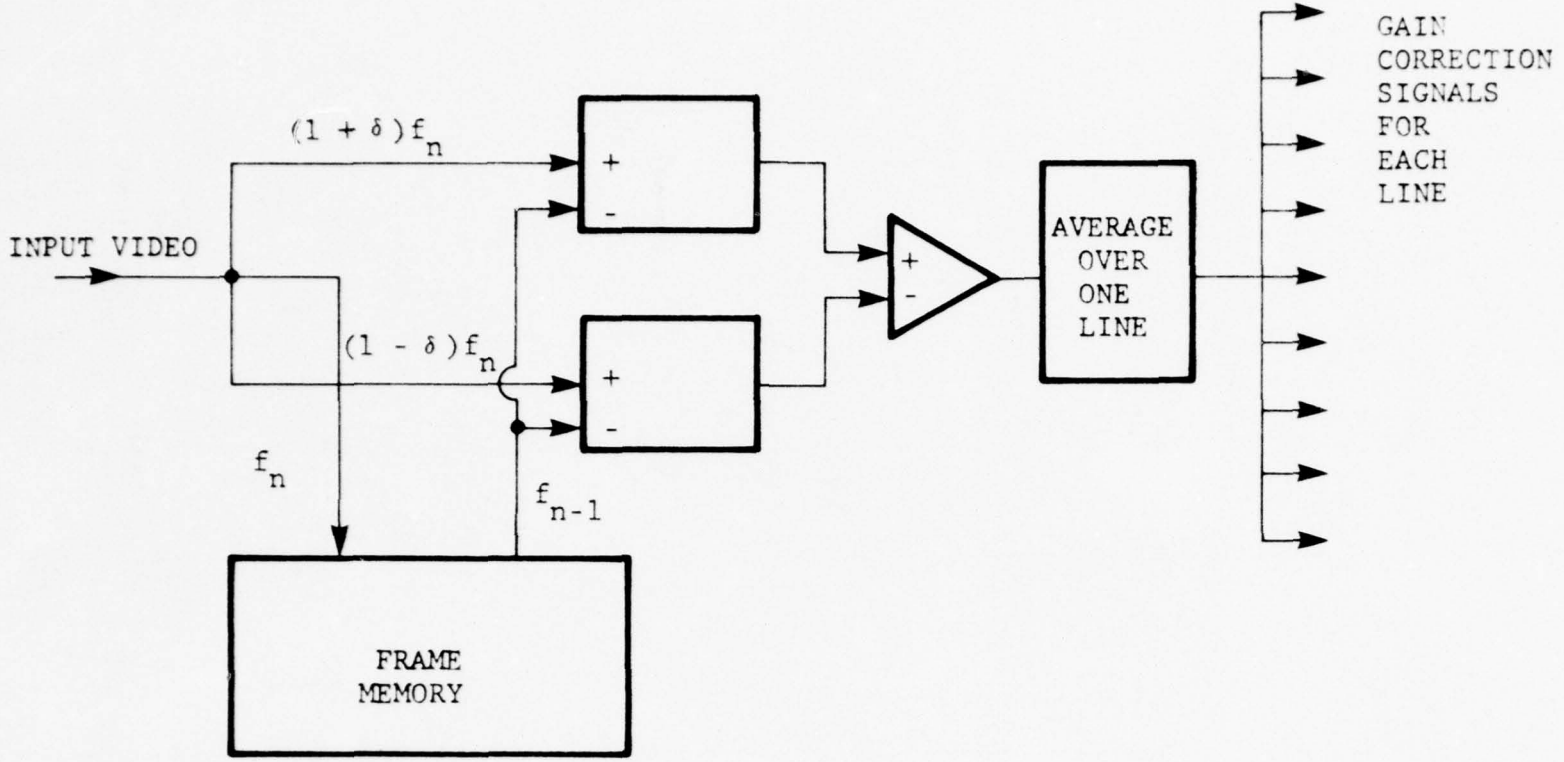
Let f_n be the current video, and f_{n-1} the prior video, after it has been brought into registration with f_n.

Consider just one line of f_n. Let be the gain error for the corresponding detector channel. That is:

$$f_{n-1} = (1 + \varepsilon)\, f_n$$

It will be shown that ε can be derived by analysis of f_n, f_{n-1} over the corresponding line. To do this the function $\emptyset$ (Figure 3.23) is formed, where:

$$\begin{aligned}
\emptyset &= \int_{\text{LINE}} \left\{ \left((1 + \delta)\, f_n - f_{n-1} \right)^2 - \left((1 - \delta)\, f_n - f_{n-1} \right)^2 \right\} dt \\
&= \int_{\text{LINE}} f_n{}^2 \left\{ \left[(1 + \delta) - (1 + \varepsilon) \right]^2 - \left[(1 - \delta) - (1 + \varepsilon) \right]^2 \right\} dt \\
&= \int_{\text{LINE}} f_n{}^2\, 4\delta\varepsilon \\
&\sim \varepsilon
\end{aligned}$$



$$\Phi = \int_{\text{LINE}} \left\{ \left((1 + \delta) f_n - f_{n-1} \right)^2 - \left((1 - \delta) f_n - f_{n-1} \right)^2 \right\} dt$$

Figure 3.23 DETECTOR-CHANNEL GAIN-CORRECTION ALGORITHM

Accordingly the gain-correction digital-word, corresponding to this line, is to be slightly increased, or decreased, according to the magnitude and sign of $\emptyset$.

In the implementation of this technique special care would have to be given to assure the stability of the correction algorithm. In particular:

- use of minimum correction gain so that correction time constant is relatively long
- the technique adjusts the relative gain of all channels (e.g., 1st vs 3rd, 2nd vs 4th, 3rd vs 5th, etc in the case when the vertical scene motion caused the 1st line of the prior frame to be in registration with the 3rd line of the current frame). Further development of the algorithm is needed to assure that it would converge to a uniform set of gains
- Special attention must be given to the detector elements at the top and bottom of the frame, so that their gains are properly corrected.

In a similar way, an offset connection may also be developed. Let

$$f_{n-1} = f_n + \alpha$$

where α is the offset error on a particular line.

To derive α, form ψ, where:

$$\psi = \int_{\text{LINE}} \left\{ \left((f_n + \beta) - f_{n-1} \right)^2 + \left((f_n - \beta) - f_{n-1} \right)^2 \right\} dt$$

$$= \int_{\text{LINE}} \left[(\beta - \alpha)^2 + (\beta + \alpha)^2 \right] dt$$

$$= \int_{\text{LINE}} 4\beta\alpha \, dt$$

$$\sim \alpha$$

Correspondingly, the offset correction digital word is to be slightly increased, or decreased, according to the magnitude and sign of ψ. The same stability considerations, as noted above for the gain correcting algorithm, apply.

3.3.7 Auto Focus

The human observer is able to focus a FLIR image by visual analysis of the changes induced in the image as he adjusts ("dithers") the focus control. An automatic focusing system is described which will perform the same function.

An optical-range-finder type of auto focusing system, in which the object is viewed (and compared) from two different ports, is undesirable for FLIR applications because of size constraints. (The relatively large diameter of even one FLIR objective is often a problem. To have two such objectives, separated by a baseline would be generally unacceptable.) What is needed is a means of automatically analyzing the image from a standard FLIR, to optimize focus.

This can be done by (automatically) commanding a small continuous oscillation ("dither") of the position of the focus lens. The resultant video signal can be analyzed, automatically, to sense which position of the dithering-lens leads to "best" focus. The focus lens is then made to move in that direction, and to execute its dither about a new point. When exactly in focus, the dither moves the lens equally in, and out, of focus so that no bias in focus quality is detected between the two dither position. The mean lens position is then not changed.

The amplitude of the focus lens dither must be small enough so that the in-focus dither yields an image which is acceptably in-focus at all times.

The first step in the design of such a system is to define the image property, or properties, which will be used to quantify the focus quality as the focus lens is dithered in, and out, of focus. Intuitively, a simple measure of the high spatial frequency content of the video should provide the required measure of focus quality. However, as noted in Reference 3.3, this is not necessarily true.

Figure 3.24 gives the MTF of diffraction limited optics for different degrees of defocus and shows how the various spatial frequencies are affected by the degree of out of focus.

(It is very important that the auto focusing system be able to sense only a moderate degree of out-of-focus (e.g., $\Delta = 1$, or less, in Figure 3.24). When the system is exactly in focus, the focus lens dither will drive the lens out of focus, by equal amounts in both directions. This degree of out of focus must be small enough to be subjectively acceptable as "in-focus" yet large enough so that the auto focusing system can reliably sense it).

Figure 3.24 shows that a $\Delta = 1$ degree of out-of-focus does not change, at all, the upper cutoff spatial frequency. It is only when $\Delta > 2$ that the upper cutoff spatial frequency is reduced. The effect (at $\Delta = 1$) is, rather, of a more or less uniform attenuation of all spatial frequencies except near zero frequency.

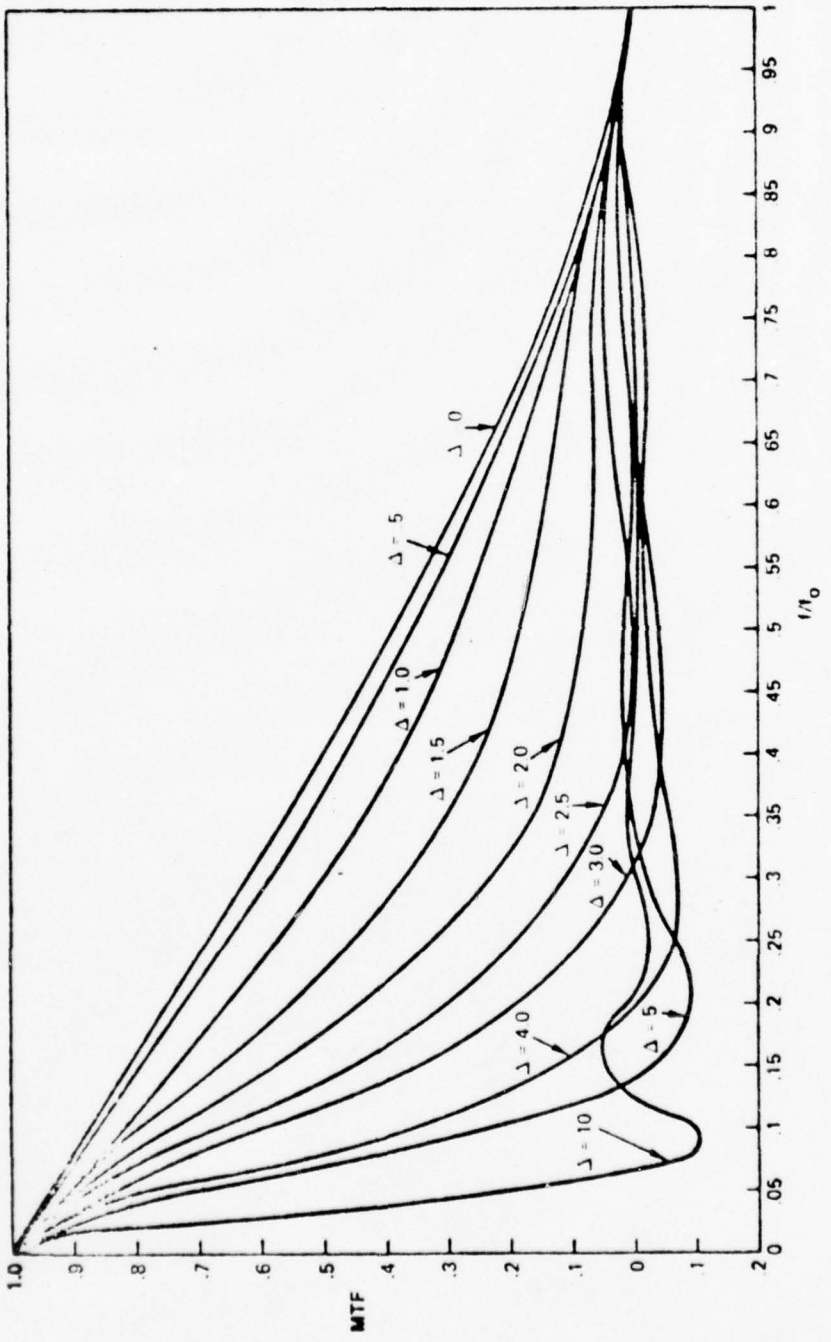


Figure 3.24 MTFs FOR DEFOCUSSED DIFFRACTION-LIMITED OPTICS
(Δ IS A MEASURE OF THE DEGREE OF OUT-OF-FOCUS,
SEE SECTION 4.6.2)

Figure 3.25 shows this more clearly. Here, the difference between the $\Delta = 1$ and the $\Delta = 0$ MTFs is expressed as an additional transfer function operating on the $\Delta = 0$ MTF. That is:

$$\delta(\text{MTF}) = (\text{MTF})_{\Delta=0} - (\text{MTF})_{\Delta=1} = \left\{ \frac{(\text{MTF})_{\Delta=0} - (\text{MTF})_{\Delta=1}}{(\text{MTF})_{\Delta=0}} \right\} * (\text{MTF})_{\Delta=0}$$

i.e., MTF loss = (transfer-function) * input MTF

It can be seen (from Figure 3.25) that the transfer function in question is (approximately) a simple band-pass filter (in two spatial dimensions).

This transfer function, expressing the effect of a moderate degree of defocus, indicates that the effect of such defocus may be more meaningfully represented as a general loss-of-contrast, rather than as a loss of any particular spatial frequency band. Moreover, since contrast can be determined by (in effect) an area correlation technique, the contrast-measure will be highly noise-free because of the large amount of spatial averaging involved. Consequently, it will be possible to restrict the amplitude (Δ) of the focus dither, thereby providing a more sharply focused image for the viewer.

Accordingly, a contrast sensing technique for auto focus will be described. To implement a contrast-sensing technique for measuring the degree of out-of-focus it is desirable that the input video by subjected to a two-dimensional band-pass filtering operation to remove those spatial frequencies that would not be modulated by a small amplitude focus dither (Figure 3.25).

The filtered video is processed, in the usual way, by the registration electronics. The objective is to generate two frames $[f(x,y)$ and $g(x,y)]$ in registration, corresponding to the two dither positions of the focus lens. (Depending on the speed of the focus lens position servo, these two frames may not necessarily be immediately successive frames.)

If the focus lens position for frame $f(x,y)$ is different to that for frame $g(x,y)$ then, in general, the contrast will also be different:

$$f(x,y) = (1 + \epsilon)\, g(x,y)$$

where ϵ represents the increase (+), or decrease (-), in contrast depending on whether the "f" focus lens position is better, or worse, than the "g' position.

The magnitude and sign of ϵ can be found by forming the function $\emptyset$ (similar to that illustrated in Figure 3.23 and described in Section 3.3.6) and evaluating over a complete, registered, frame:

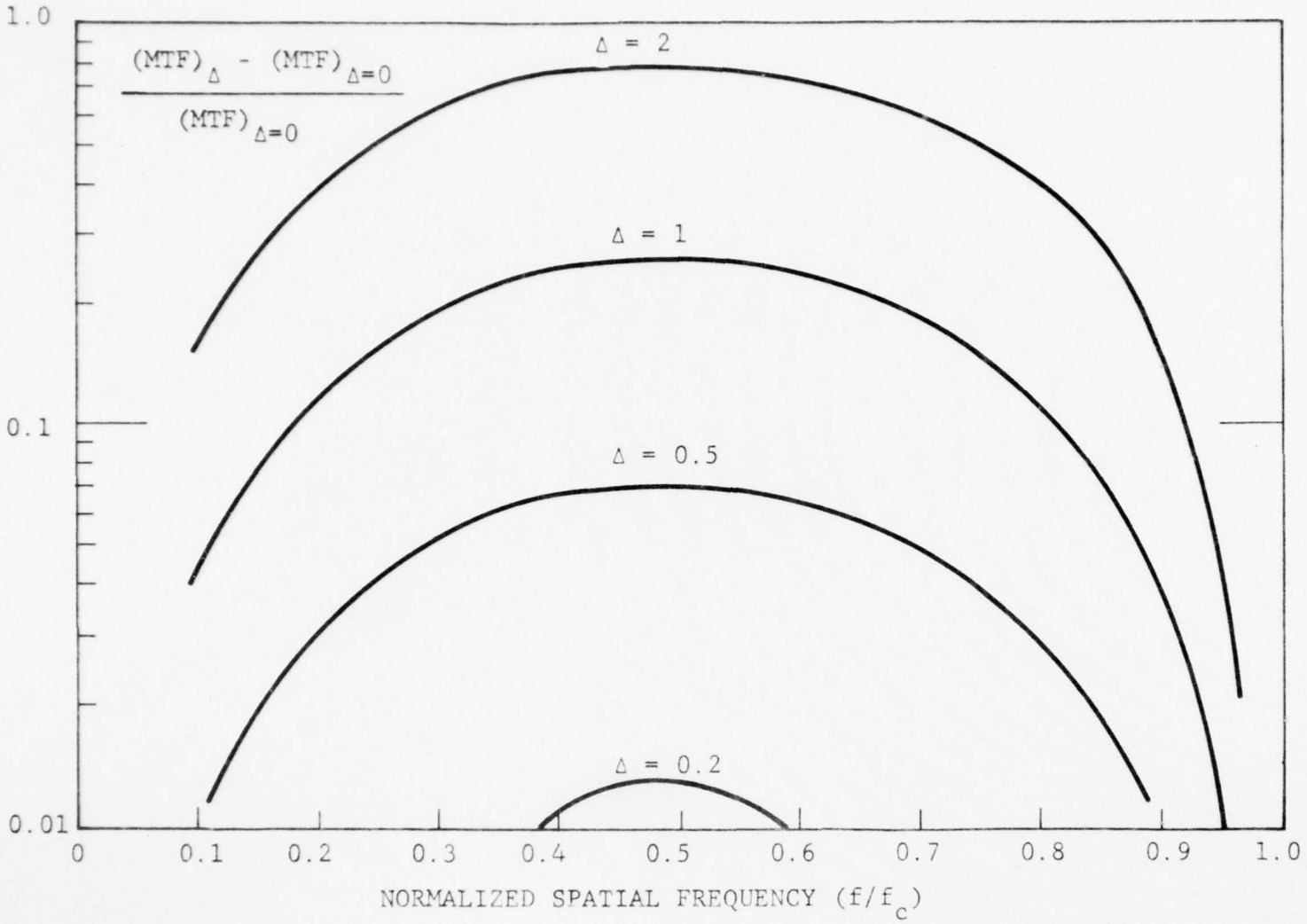


Figure 3.25 DEFOCUS TRANSFER FUNCTION (ADAPTED FROM REFERENCE 3.4)

$$\emptyset = \iint_{\text{FRAME}} \left\{ \Big((1+\delta)\, g(x,y) - f(x,y) \Big)^2 - \Big((1-\delta)\, g(x,y) - f(x,y) \Big)^2 \right\} dxdy$$

$$= \iint_{\text{FRAME}} g^2(x,y) \left\{ \Big((1+\delta) - (1+\varepsilon) \Big)^2 - \Big((1-\delta) + (1+\varepsilon) \Big)^2 \right\} dxdy$$

$$= \iint_{\text{FRAME}} g^2\,(x,y)\; 4\delta\varepsilon\, dxdy$$

$$= 4\delta\varepsilon \iint_{\text{FRAME}} g^2\,(x,y)\, dxdy$$

The function ψ will also be derived:

$$\psi = \iint_{\text{FRAME}} g^2\,(x,y)\, dxdy$$

Then, the value of ε can be determined independently of scene content:

$$\varepsilon = \frac{\emptyset}{4\delta\psi}$$

Since ε is normalized (independent of scene content) there is a direct (constant) relationship between the magnitude and sign of ε and the change in the focus lens position required to bring the system into focus. Thus ε can be applied directly to the focus servo. (If the scene content is very small, then ψ will be small, and the technique will obviously break down. A threshold value of ψ should be set, below which the auto focus loop should be de-activated.)

An important advantage of this area-contrast focus sensing technique is that it is immune to variations in scene content. A focus technique based, for example, upon sensing the variations in the high spatial frequency content of the image may(as noted in Reference 3.3) falsely interpret changes in the spatial frequency content of the (moving) scene itself, with changes induced by the focus-lens dither. It can be seen from Figure 3.25 that the latter must necessarily be very small (a few percent) since the focus lens dither must be small enough so that the image is always acceptably in focus. Accordingly, the variations in the spatial frequency content, due to the scene moving through the frame, should be kept to the order of 1% or less over the dither cycle to avoid disturbance of the focus sensing system. This implies a requirement for highly accurate scene tracking in a focus system that is based upon sensing high frequency detail. This tracking accuracy is inherent in the registration/contrast-sensing system.

In some situations it may be desired to focus upon only one (small) part of the video frame - for example when the small detail of interest is at a different range than the general background and both cannot be in focus at the same time. This can be done by operator designation of a video tracking gate around the detail of interest. In this case, the functions $\emptyset$ and ψ are evaluated only inside the gate region. Once designated, the gate can be made to automatically follow (track) the detail of interest by utilizing the 8 scene motion parameters that are derived as part of the image registration operation.

When the auto-focus action is restricted to a small part of the video frame, the requirement for accurate scene tracking, in a system based upon sensing changes in the high frequency content induced by lens dither, becomes even more stringent.

3.3.8 Image Stabilization

The image registration technique that has been described so far is applicable, generally, to moving video imagery when the image movement is smooth and relatively predictable. That is, the change in the nature of the scene motion between successive frames is small - even though the motion itself may be large. This is the kind of motion induced by linear and/or angular velocity of the sensor relative to the scene as illustrated in Figure 3.2.

Image motion caused by vibration is of different type. In this case the motion itself may be relatively small, but will change rapidly from frame to frame.

The image registration electronics can be applied to eliminate either, or both, types of motion in the displayed image.

In the normal utilization of image-registration, the current frame is always displayed, full-frame, as received from the sensor. The prior frame(s) are displaced by an address-modified memory-read operation so as to be in registration with the current frame.

However, by address-modifying the memory "write" operation, instead of the "read" operation, the display can be "frozen" at some specified prior frame. All subsequent frames are then written, and displayed, relative to the initial scene structure and will not, therefore, appear to move.

To stabilize against image motion of the second type (vibration), inertial sensors can be attached to the FLIR to detect the angular accelerations of the FLIR that (alone) cause this type of image motion. The inertially-sensed image motion information can then be used to augment the motion information detected by analysis of successive video frames. The inertial information is inherently much faster (higher bandwidth) than the picture-processed information.

The high frequency (inertial) and the low frequency (picture), motion information can be used to address-modify (either in the write, or in the read, operation) so as to eliminate the effects of scene motion and vibration in the displayed image.

3.3.9 Moving Target Indication

In this application the operator may choose to enhance all moving targets within the scene and suppress background.

As the sensor is flown over the target scene, the "static" background and the moving targets would normally both appear to move in the video display. The differential motion of small parts of the scene relative to the large area background (e.g., a truck moving along a road) is enhanced in this MTI mode.

To provide the MTI feature successive registered frames are subtracted. "Static" background -- being in registration in the two frames -- is thereby cancelled and the small moving targets enhanced.

The operator may place a small tracking gate over a moving target so that the system can derive two independent interframe displacement functions: one for the large region outside of the gate, the other for the small region within the gate. In this way, the system can not only detect moving targets but also display them more clearly to the operator through integration, in registration, of successive target images.

3.3.10 Continuous Display with Intermittent Video

In some situations the operator may see only intermittent scene images on his display. This may occur, for example, when looking through thin clouds, or during temporary RPV data-link loss (due to jamming or line-of-sight obstructions).

Intermittent views of the target scene may be very distracting to the operator, and limit his search effectiveness.

The image registration electronics can be applied to provide the operator with a continuous uninterrupted view of the target scene under these conditions.

There should be appreciable scene overlap (e.g., 50%) between the last frame before loss-of-picture and the first frame after the break. The synthetic picture generated during the break should be clearly flagged, as such, to the operator. For example, it might be made to flicker.

The loss of picture may be immediately sensed either as a loss of RF carrier (in the case of an RPV), or by the detection of sudden large differences on comparing successive frames (for example, as clouds flash into the picture).

Upon detection of a loss-of-picture condition the operator's display is fed with the last frame stored in digital memory, continuously address-modified according to the last measurement of the scene motion parameters, so as to simulate a continuing scene motion. Thus the operator continues to see natural moving imagery without any apparent interruption. No new scene information will, of course, be displayed while the sensor video is temporarily lost but the advantages to the operator of the dynamic display provided from the image-memory are:

1. He is not distracted by the transient loss-of-picture condition

2. He can spend his time usefully, during the temporary loss-of-picture condition, searching for targets in the "old" imagery or, in a pilotage application (not involving ground obstacles) continuing to fly the aircraft

3. When the true (sensor) picture is restored he is not distracted by this transient, and does not have to reacquire the target and/or his frame of reference.

3.3.11 Zoom

Fast target acquisition/recognition/strike, in daylight and at night, is an important requirement of air-ground weapon systems for both manned aircraft and RPVs. Because of the inherent field-of-view/resolution dilemma, the operator of a thermal imaging system must, except at short ranges, change the system field-of-view as he progresses through the task sequence: search-acquisition-recognition-aim point designation. Wide field-of-view, relatively low resolution, is required for acquisition: narrow field-of-view, relatively high resolution, is required for recognition.

Conventionally a field-of-view change is accomplished by lens switching. Two, or more, lenses of different focal lengths are mounted on a lens turret. By rotating the turret, a new lens is brought into position, thereby changing the sensor field of view. On acquiring a target in his wide field-of-view display, the operator re-aims his sensor so as to bring the target to the center of his display, and then commands a lens change to obtain a narrow-field, higher resolution, view of the target. The field-of-view change causes a large, abrupt, change in the image displayed on the operator's screen.

In the context of quick-reaction air-ground missions (in some cases involving a single-place aircraft) this conventional method of changing system field-of-view is seen as a potentially disorienting and high workload task, adversely effecting the speed and reliability of target acquisition/recognition. The specific problems are:

- The need to perform a sensor aiming task to bring the target to the center of the wide field of view
- The disorientation induced by the abrupt change in the displayed image after changing field of view
- The need to reacquire the target in the close-up view
- Uncertainty that the target acquired in the wide field-of-view is actually present in the narrow field display
- Disturbance to the pilotage function caused by the disorienting field-of-view changes.

The reason that field-of-view change is effected (now) by lens-switching, in spite of the problems noted, is (at least in part) because of the practical difficulties and optical limitations of zoom lenses.

3.3.11.1 Zoom Effect Using Fixed Lenses

The registration electronics can be applied to provide a continuous zoom effect, electronically, without loss of resolution, using the existing sensor with its conventional fixed (not zoom) lenses.

Just before initiating a lens change, the current image, and its rate and type of motion, are captured in the memory system in the usual way (Figure 3.26). On executing the lens change, this circulating image data is zoomed by introducing (independently of the moving image processor) the appropriate address modification. This continuously zooming image is displayed to the operator. As soon as the image from the next lens selection becomes available, it is appropriately injected into the zooming image data, then circulating through the memory. The zoom action continues until, finally, only the narrow-field video is presented to the operator.

During the zoom action there is some loss of resolution, but only in the edges of the picture where the last (captured) frame from the wide angle sensor is zoomed. There is no loss of resolution in the narrow angle sensor video as it is zoomed up to full display size. As the magnitude of the resolution loss increases, the peripheral area over which it occurs decreases. In the terminal states of the zoom there is, of course, no loss of resolution. There is never any loss of resolution in the foveally-viewed region of the picture. Since the zoom, by the very nature of the proposed application, will be very fast (approximately 1 s), the temporary peripheral resolution loss during the zoom will not be of practical significance. (The static acuity of peripheral vision is low. The dynamic peripheral acuity in viewing a rapid zoom is even lower because of retinal smearing (due to image motion) in the peripheral areas of the rapidly zooming picture. In other words the observer will be relatively insensitive to any temporary reduction of peripheral resolution that may occur in the display during the proposed zoom).

3.3.12 Signal Processing for Pyroelectric Vidicon

The failure of the pyroelectric vidicon (or pyroelectric array) to respond to steady image requires that it be operated in either a panning, orbiting or chopping mode. The chopping mode (as described by Helmick and Woodworth of NWC at the 1975 IEEE Symposium on Applications of Ferroelectrics) provides the best imagery. However, it requires a single frame of video storage. NWC used a video disc for this purpose.

The registration electronics can be adapted to provide the necessary single-frame storage and also, with the same memory, to integrate the chop-demodulated video to provide improved signal-to-noise ratio.

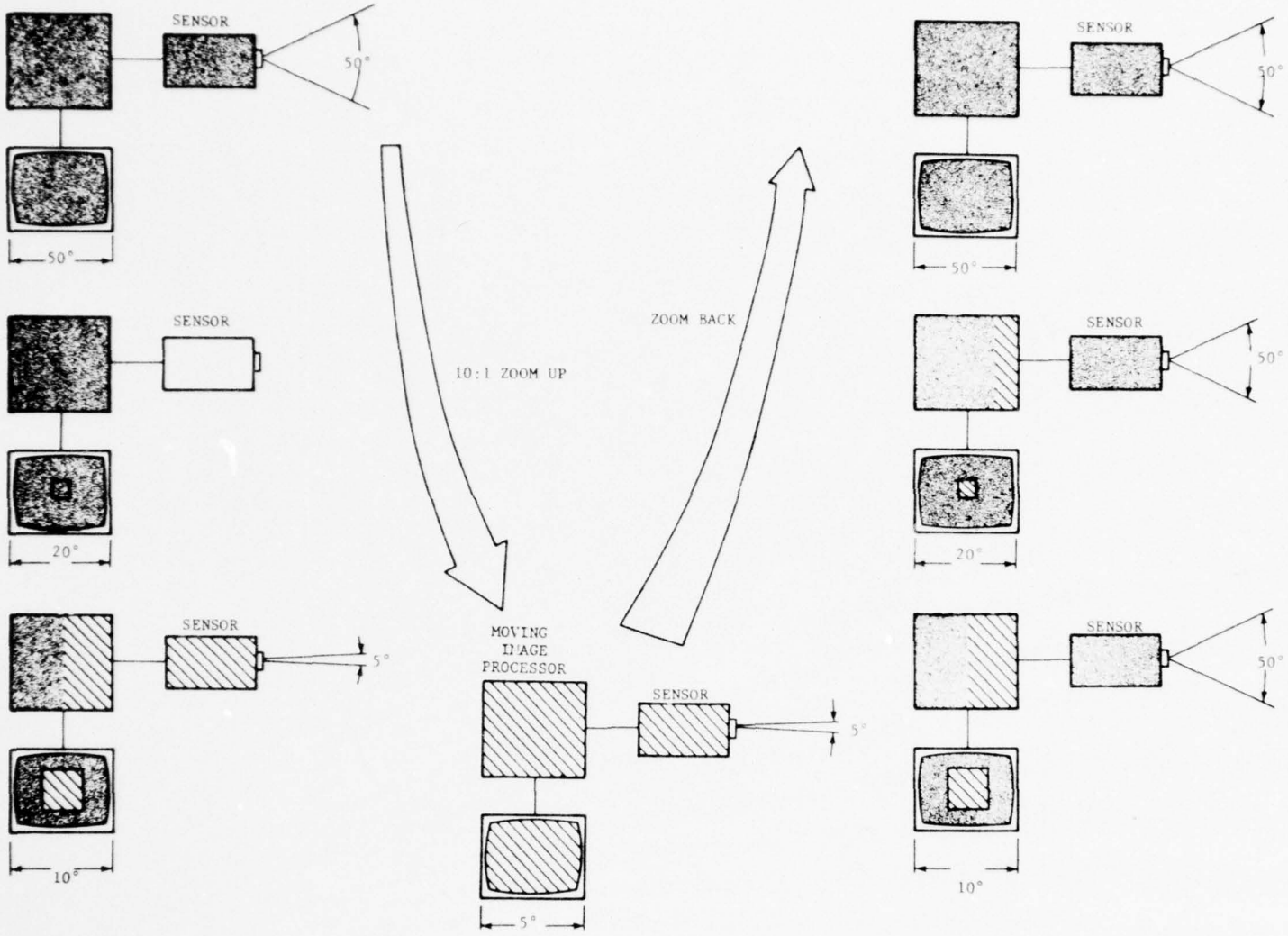


Figure 3.26 CONTINUOUS ZOOM DISPLAY USING FIXED LENSES

The large improvement in NETD resulting from integration would be particularly useful for the pyroelectric vidicon because of its relatively low thermal sensitivity. It may be possible to exploit the large improvement in NETD by using smaller, less expensive, lenses than the F/1 of F/0.7 normally used with a pyroelectric vidicon.

For pyroelectric processing, the video signal (I') plus pedestal (G) in a frame when the vidicon faceplate is irradiated by the scene (chopper wheel "open") is added to the signal read out from memory (V_m) according to the law:

$$V_m' = V_m (1 - 1/G) + P + I'$$

For scene detail appearing in the video frame for the first time, the law is:

$$V_m' = G (P + I')$$

The displayed signal (pedestal-free) is V_m (no display of first-time detail).

On the next frame the chopper wheel is in the "closed" position, and the video consists of the pedestal (P) with a negative signal (-I'). Accordingly, the add-to-memory logic for this frame is:

$$\begin{aligned} V_m'' &= V_m' - (P - I') \\ &= V_m (1 - 1/G) + P + I' - (P - I') \\ &= V_m (1 - 1/G) + 2I' \end{aligned}$$

For first-time detail, the law is:

$$V_m'' = V_m' - G (P - I')$$

$$V_m'' = 2GI'$$

On this frame, the displayed signal is V_m'', also pedestal-free.

The total effect is:

1. only the true scene-signal, free of pedestal and vidicon defects, is integrated in the memory;
2. the pedestal is never displayed;
3. the usual benefits of enhanced sensitivity through integration apply;
4. a normal video signal is generated with moving, and with static, scenes;
5. an integrated signal is available for display on all frames, including those with chopper blade in the "closed" position;
6. no moving parts image-memory.

3.3.13 Modularity

Thermal imaging devices are used in many different applications. It would be prohibitively expensive to develop a new FLIR system for each application, yet it is impossible, because of the diverse nature of the requirements, to optimally satisfy the various requirements with a single system.

This had led to the concept of modularity, in which any particular FLIR system would be constructed from a relatively small number of more or less standard modules ("optics", "focal plane", "scanner", "electronics", etc). In this way the economy of standardization may be achieved, while providing the required range of performance characteristics.

The registration electronics may prove to be a useful module, within this concept.

An immediate example is the ability provided of flexible tradeoffs between data-rate and sensitivity, while using the same scanner and scan rates.

Moreover, with relatively little modification in each case, the basic (common) image registration electronics can provide a broad range of features (see Figure 3.7). Thus the image-registration electronics, itself, may be modularized, so that any desired set of features can be economically provided.

REFERENCES

3.1 Moving Image Processor, October 1976, Honeywell Radiation Center.

3.2 Digital Operations for Image Registration, October 1976, Honeywell Radiation Center.

3.3 Final Technical Report for the Video Controlled FLIR Focus Program, Texas Instruments Incorporated, August 1976, AFAL-TR-76-161.

3.4 Tables of Modulation Transfer Function of a Defocussed Perfect Lens. L. Levi and R.H. Austing. Applied Optics, 7, pp 967-974, May 1968.

3.5 Thermal Imaging Sensors, J.M. Lloyd, Plenum Press, 1975 (page 132).

SECTION 4

AUTO-FOCUS

4.1 INTRODUCTION

The task of manually focusing a FLIR represents additional workload for an operator who may already be heavily stressed. Accordingly, consideration should be given to the provision of some means of automatic focus control.

Before considering possible auto-focus techniques, it will be useful to review the factors that can cause a FLIR to be out of focus. It will be seen that there are several such factors whose relative significance depends upon the particular circumstances in each case.

4.2 FACTORS THAT CAUSE DEFOCUS

4.2.1 Range Effects

Range effects (that is, the spread of possible object ranges relative to the depth of field of the FLIR) are often considered as the main (if not the only) reason why a FLIR must be focused.

In fact, range effects may not be the main problem. For example, the largest objective lens diameter of a FLIR is typically of the order of 8 inches. If such a FLIR is focused for infinity it is evident that the defocus blur circle diameter at the object will be 8 inches at <u>all</u> ranges, from zero to infinity. By focusing the system at some intermediate range the defocus blur circle diameter can be kept within 4 inches over a wide spread of object ranges. For many target-detection tasks these focus blur magnitudes would be of little practical significance. That is, the system could be operated, quite adequately, with fixed focus. In this case, the operator would have no focusing task at all.

Range effects cannot be totally ignored, however. Their practical significance depends upon the FLIR and on the mission. Accordingly, range-effects are analyzed in detail in Section 4.6 in terms of the focusing difficulty that they induce. A quantitative measure of focusing difficulty is proposed and it is shown how this may be estimated for any particular set of conditions.

4.2.2 Temperature Effects

Variations in the temperature of the objectives lens (and, to a lesser extent, of the entire optics framework) can produce significant defocus effects. These effects are analyzed in Section 4.5. It is shown that the degree of defocus is proportional to $(D\Delta T/F\lambda)$ where ΔT is the temperature "error" of the lens, F is its F# and D its diameter. For a 0.25 m objective at 10μm, temperature variations of the order of 1^{o}, or more, can produce significant defocus.

4.2.3 Mechanical Effects

An axial displacement of the scene image, relative to the FLIR focal plane, of only a few thousandths of an inch may cause significant defocus. Having regard to the relatively large optics, the mechanical scan system, the lens changing mechanisms, etc., that are typical of FLIRs, it is evident that defocus effects may occur due to miscellaneous mechanical effects.

4.2.4 Operator Effect

If the operator is provided with a wide-range focus control (so that he can correct for any, or all, defocus effects) he may, himself, contribute the largest defocus effect. When attempting to focus a very low contrast scene the operator will (typically) swing his control from one extreme to the other. Most of the time the system will then be totally out-of-focus and (with a low contrast scene) no scene image will be detectable. If this focus search is done too rapidly, the operator may miss the brief "flash" of the focused scene image that will occur as the focus control is moved through the sharp-focus position.

These same considerations apply, also, to a "true" auto-focus system. That is, the speed and/or reliability of an auto-focus system may be adversely affected if the system must search (over a wide focus-range) for optimum focus when the scene image has very low contrast.

4.3 FOCUS CONTROL SYSTEMS

Three types of focus control systems can be identified:

(1) "true" auto focus - an electronic system analyses the FLIR video and controls the focus lens position so as to (automatically) cause the scene image to be sharply focused.

(2) temperature compensation - the temperature of the objective lens is continuously sensed and this information is used to appropriately compensate the focus lens servo.

(3) auto-collimation - a test image at infinity (or at some other specified range) is injected into the FLIR; the FLIR is then manually (or automatically) focused for this test image.

A "true" auto-focus system is described in Section 3 . It exploits the fact that the contrast in the image is maximized at sharp-focus. An important advantage of contrast-sensing is that it can yield a scene-independent measure of focus-state. A "true" auto-focus system compensates for <u>all</u> defocus effects, including range effects. The temperature sensing and auto collimation techniques do not compensate for range-effects. In fact, in the recommended form of these latter systems, the operator would be provided with a manual focus control calibrated in range. The purpose of the temperature-sensing, or auto-collimation, technique is then to guarantee that the FLIR will, in fact, be focused for whatever range the operator may "dial-in" on his focus control. If adequate* range information is available (either from the operator, from a range

*When the focusing difficulty (as defined in Section 4.6) is small (as it will be, in most cases) only a very crude estimate of range is needed to give sharp focus.

finder, or from a navigational system) this technique might be preferable to a "true" auto-focus system since it avoids the possibility of wide-amplitude focus-search which, with low-contrast scenes, may result in unreliable or lengthy acquisition. Moreover, for applications in which range-effects are minimal, there is no specific need for true auto-focus.

The temperature sensing technique has the advantage of simplicity but is "open-loop". That is, it does not depend on sensing focus-quality, so that it could be disturbed by mechanical errors, temperature sensing errors, temperature gradients, etc.

In the auto-collimation technique, a test image (at a known virtual range) is presented to the FLIR, and the FLIR is then manually (or automatically) focused for this image.

The advantage of the auto-collimating technique (over a "true" auto-focus system) is that it can allow the FLIR to be instantly, and sharply, focused even in the complete absence of any scene detail. This could be very important for long range target acquisition - for examples in an air-ground system in which the initial (acquisition) slant range may be up to 20,000 ft and survivability may depend, critically, on fast, reliable, target acquisition.

At such long ranges, under low contrast conditions, there may be little or no scene detail in the FLIR video with which a "true" auto-focus system could function. Moreover, at long ranges there is generally no need for range-correction to focus. That is, the FLIR should, simply be focused for infinity. It is evident that under such conditions an auto-collimating system would be highly effective (providing instant sharp focus) whereas a "true" auto-focus system might be slow, unreliable or even totally inoperative.

The principle behind the auto-collimating technique is that when a plane mirror is placed in front of the objective, the operator will see a sharply focused image of the cold focal plane when, and only when, the FLIR is focused for infinity (Figure 4.1). The plane mirror reflects "cold" radiation from the focal plane back onto itself (narcissus effects). The range-accuracy of this technique depends upon the curvature of the "plane" mirror - which can be easily set, and maintained, at nominally zero curvature.

The technique could be implemented in a number of ways. For example:

(1)	Focal plane	(a)	unmodified (utilize pattern detail)
		(b)	special pattern on the focal plane
(2)	Mirror Size	(a)	Same as objective
		(b)	Smaller than objective
(3)	Mirror Reflectivity	(a)	100 percent
		(b)	<100 percent (beamsplitter)
(4)	Mirror Position	(a)	in place all the time
		(b)	snapped out of position when not in use

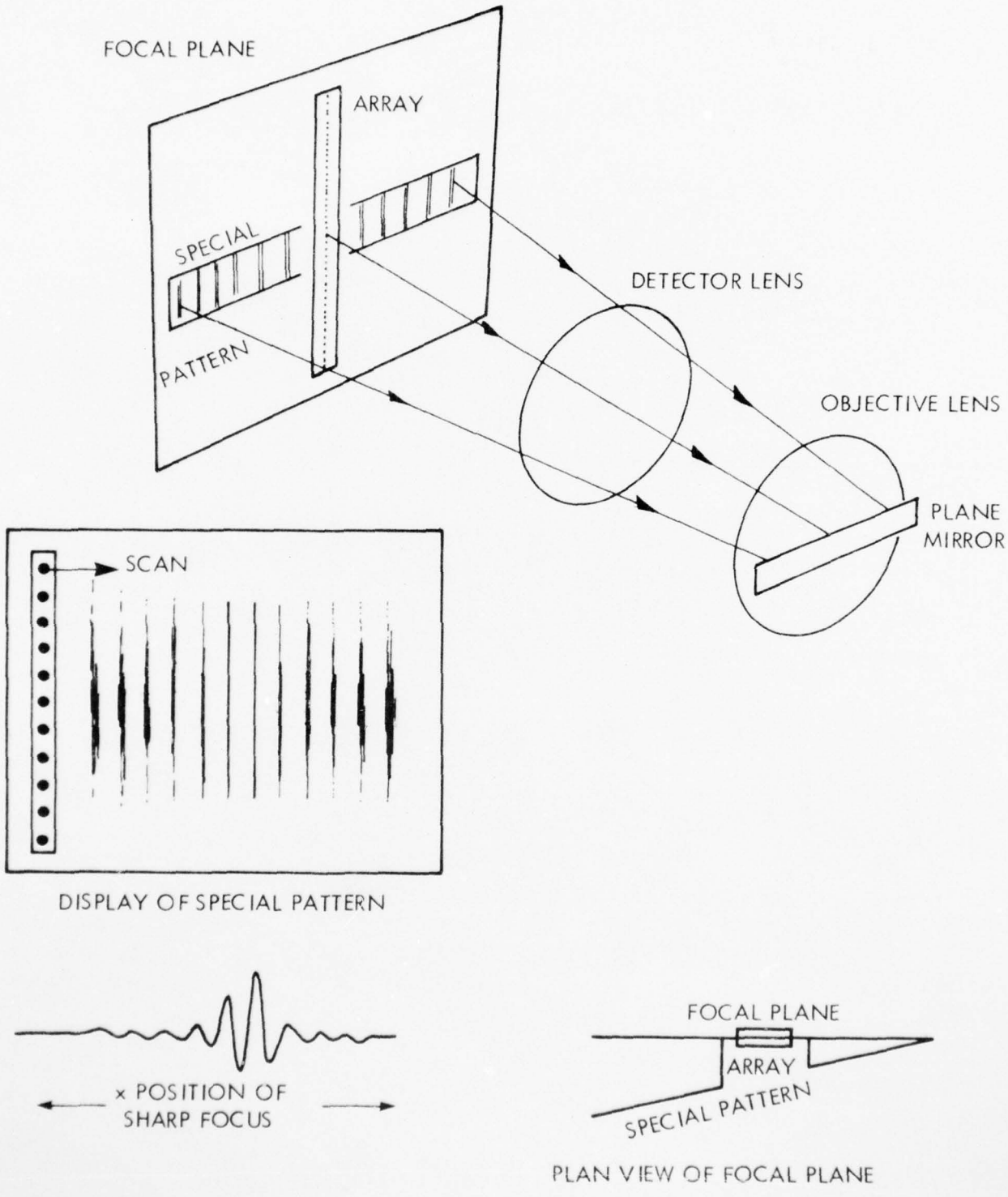


Figure 4.1 AUTO-COLLIMATION FOCUS TECHNIQUE

One particular implementation is illustrated in Figure 4.1. A special pattern is formed at the focal plane consisting of alternating vertical stripes of black and reflecting material. The black material is seen (by the mirror) as relatively cold (focal plane temperature), whereas the reflecting stripes are seen as relatively warm (average scene temperature). The plane mirror is rectangular, covering the full diameter of the objective in one dimension and only a small fraction of the objective diameter in the other direction. In this way it provides negligible obscuration to the input rays. The relative orientations of the plane mirror and of the special focal plane pattern are such that a sharp image is formed (by reflection at the plane mirror) of the high spatial frequency in the pattern, but the image of the special pattern is heavily smeared-out (by diffraction) in the other dimension.

The detection of the diffracted pattern could be done either by eye or automatically. In the latter case the video would be specially filtered (in both x and y) to enhance the narrow band of spatial frequencies carrying the focus information. (A similar filtering action is performed by the eye, in the case of visual detection.)

The focal plane pattern represents a relatively large amplitude temperature sinusoid. The plane mirror, because of its small area, reflects only a small fraction of the pattern energy back onto the focal plane. The detected image of the pattern, encompassing only a narrow range of spatial frequencies, can be enhanced by filtering. These three factors provide a range of design flexibility which can be utilized to generate an easily detectable focus signal.

By arranging for the focal plane pattern to be formed at a slight angle to the focal plane (Figure 4.1) the FLIR cannot only be focused for infinity but also for any specific intermediate range. To illustrate this, suppose that a lens is to be focused for a range R. Then, the focal plane must be located at a distance $f + \epsilon$ from the lens (focal length f) where

$$\frac{1}{f+\epsilon} + \frac{1}{R} = \frac{1}{f}$$

i.e.,

$$R \approx f^2/\epsilon$$

Suppose, now, that the special focal-plane pattern is located at distance $f - \epsilon$ from the lens. Its image will appear (to the mirror) to be at a range S where

$$\frac{1}{f+\epsilon} + \frac{1}{S} = \frac{1}{f}$$

i.e.,

$$S \approx -\frac{f^2}{\epsilon}$$

However, after reflection in the plane mirror the image S', of S, will appear at a range

$$S' = -S = \frac{f^2}{\epsilon}$$

In other words, with a slight axial displacement of the special pattern from the focal plane, the system auto-collimates to a finite range that is proportional to the reciprocal of that displacement.

In the system illustrated in Figure 4.1, the axial displacement of any part of the pattern is proportional to its "x" position. Thus the range at which the FLIR is focused is proportional to the position along the x axis of the most sharply focused section of the focal plane pattern. In other words the x axis of the display can be directly calibrated in range. To focus the FLIR for any particular range, the operator would adjust his focus control so as to cause the display of the special pattern to be sharply focused at the appropriate x position.

The special focusing pattern appearing on the operator's display would also be useful as a visual system-performance check. However, in normal operation it should be removed from the display. This could be done by rotating the rectangular mirror so that it did not then reflect the pattern back to the focal plane. In other words the auto-collimating system would be used only occasionally - to "zero" any errors in the range calibration of the manual focus control.

In an automatic mode of operation, the pattern would always be present in the video but would be removed from the displayed video by means of a suitable band-stop filter. This filter would, of course, also remove the corresponding narrow-band of spatial frequencies from the scene detail - but the effect of this may be insignificant.

4.4 RECOMMENDATIONS

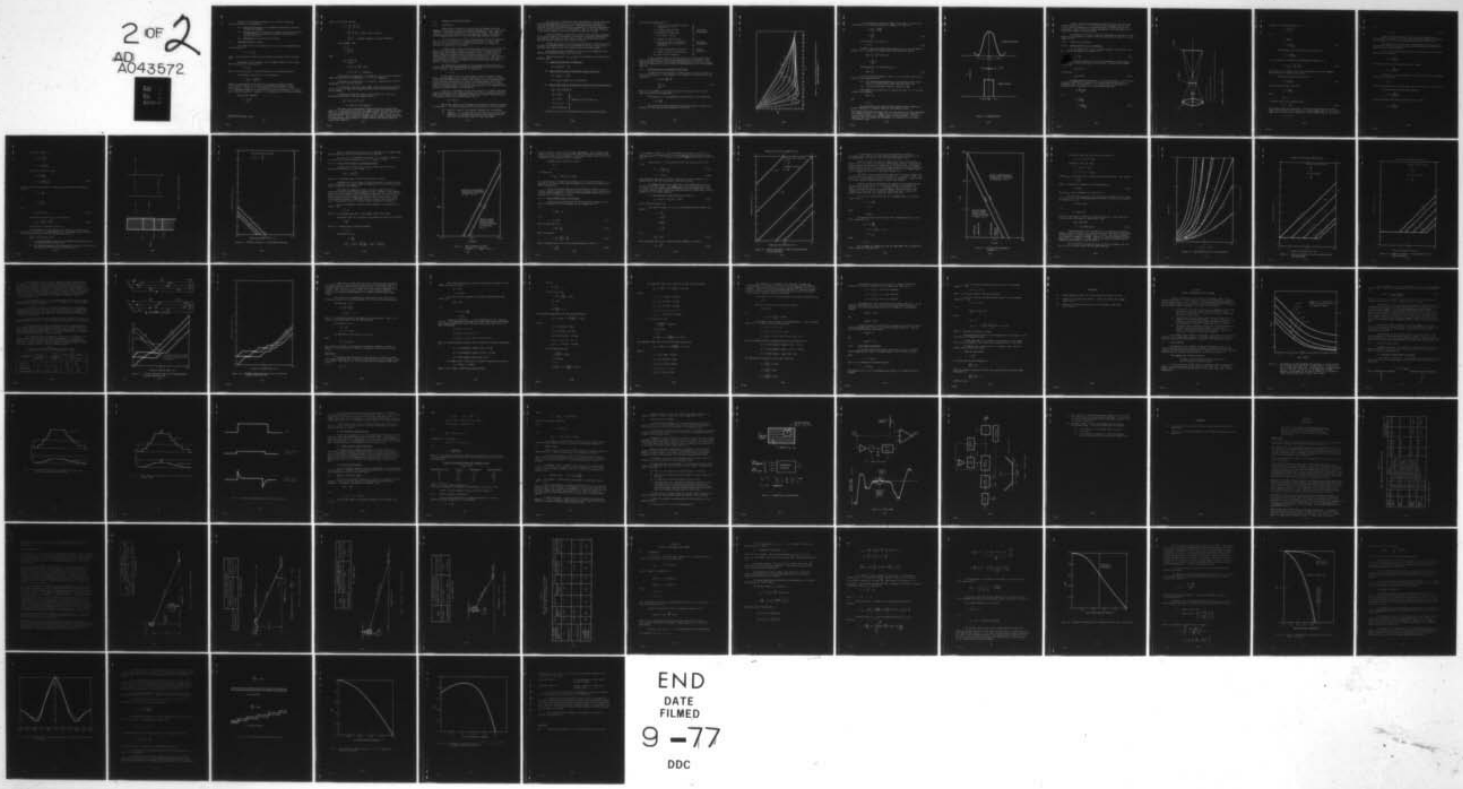
For applications in which range-effects (on focusing-difficulty) are minimal, consideration can be given to the use of either the temperature-compensation, or of the auto-collimation, techniques. Both techniques have the potential for very simple implementation. The auto-collimation technique offers the additional operator benefit of a visual system-check feature. (If he can see the special pattern at its usual clarity the operator can be assured that the FLIR is operating correctly.) By eliminating the need for large-amplitude focus correction (to compensate for temperature effects on focus) the operator's focusing task is thereby reduced to just compensation for range-effects. (His focus control should be calibrated in range - in many cases the operator need 'dial-in' only a very approximate estimate of range to achieve adequate focus.) In this way the operator should be able to acquire a low-contrast target directly; without the wide-amplitude focus search procedures which can result in unreliable and slow acquisition.

For applications in which the focusing difficulty is due, in large part, to range effects (and suitable range information cannot be made available) then a true auto-focus system is required. A combination of true-auto-focus with one of the other two techniques would improve focus-acquisition (speed and reliability) if rapid (focus) acquisition proved to be a problem with low contrast imagery.

Finally, if a low cost, simple, and reliable true auto-focus system could be developed it might be used for all applications, whatever the focusing effect(s) might be.

AD-A043 572 HONEYWELL RADIATION CENTER LEXINGTON MASS F/G 17/5
AUTOMATED IMAGE ENHANCEMENT TECHNIQUES FOR SECOND GENERATION FL--ETC(U)
AUG 77 DAAG53-76-C-0195
UNCLASSIFIED HRC-77-5-1 NL

2 OF 2
AD A043572



END
DATE FILMED
9-77
DDC

Research and development (leading to one, or more, operational focusing systems) is recommended for:

(1) Practical Evaluation of the temperature-compensation and auto-collimation techniques.

(2) Operator Tests with an accurately range-calibrated focus control to determine when range-effects contribute significantly to operator work load.

(3) Development of a true auto-focus system.

4.5 TEMPERATURE EFFECTS ON FOCUS

The change (Δf) in focal length (f) of a lens due to temperature error (ΔT) is given by:

$$\Delta f = - f \Delta T \frac{\partial \eta}{\partial T}$$

where η is the refractive index of the lens material (equation 6.98, Reference 4.1).

The change in focal distance, due to thermal expansion of the framework holding the optics is given by:

$$\Delta f = f \alpha \Delta T$$

where α is the coefficient of thermal expansion of the framework material.

As discussed in Section 4.6.2 , the parameter Δ,

$$\Delta = 2\left[\sin^2 \left(\frac{D}{2f}\right)\right] \frac{\Delta f}{\lambda},$$

(where D is the diameter of the lens) is a convenient normalized measure of the degree of out-of-focus of an optical system corresponding to the focal error Δf. A value of $\Delta = 0$ corresponds to perfect focus. A value of $\Delta = 1$ is generally taken as the limit of focus error corresponding to an acceptable "in-focus" condition.* Values of $\Delta > 1$ corresponds to out-of-focus conditions.

From the last equation

$$\Delta \approx \frac{1}{2F^2} \frac{\Delta f}{\lambda}$$

*See footnote on page 4-13)

where F is the F# of the lens

$$\therefore \Delta = \frac{1}{2f^2} \frac{f\Delta T}{\lambda} \frac{\partial \eta}{\partial T}$$

$$= \frac{D}{2f} \frac{\Delta T}{\lambda} \frac{\partial \eta}{\partial T} \quad \text{(thermal effect on lens)}$$

$$= \frac{D}{2f} \frac{\Delta T}{\lambda} \alpha \quad \text{(thermal expansion of optics framework)}$$

As an example, take

$F = 2$

$D = 0.25$ m

$\lambda = 10^{-5}$ m

Then,

$$\frac{\Delta}{\Delta T} = \frac{0.25}{4 \times 10^{-5}} \frac{\partial \eta}{\partial T}$$

$$= 6.25 \times 10^{-3} \frac{\partial \eta}{\partial T} \quad \text{(lens)}$$

$$= 6.25 \times 10^{-3} \alpha \quad \text{(framework)}$$

From Table 6.1 in Reference 4.1, values of $\partial\eta/\partial T$ for typical materials range from $300 \times 10^{-6}/^{o}C$ (germanium) to $51 \times 10^{-6}/^{o}C$ (IRTRAN 2).

From Table 6.3 in Reference 4.1, values of α range from $26 \times 10^{-6}/^{o}C$ (magnesium) to $5.7 \times 10^{-6}/^{o}C$ (Ni 42).

It is evident, therefore, that thermal effects on the lens itself are of the order of ten times more significant than on the framework holding the optical system.

Assuming an intermediate value of $150 \times 10^{-6}/^{o}C$ for $\partial\eta/\partial T$, the sensitivity (on focus) of the lens-thermal effect is:

$$\frac{\Delta}{\Delta T} = 6.25 \times 10^{-3} \times 150$$

$$\approx 1 \text{ (unit of } \Delta \text{ per degree C)}$$

This means that uncompensated temperature variations greater than about $1^{o}C$, in the objective lens (in the example system) would defocus the image. Similarly, uncompensated temperature variations greater than about 10^{o} in the optics framework would also defocus the image. Both of these effects are proportional to D/λ. Lower resolution systems (than the example system) would thus be correspondingly less sensitive to thermal effects, and higher resolution systems more sensitive.

4.6 ANALYSIS OF FOCUSING DIFFICULTY

4.6.1 Introduction

Under certain conditions an electro-optical imaging system (for example, a FLIR) may not require any operator-adjustment of focus. Under other conditions, continuous adjustment of focus may be necessary. The degree of difficulty of the focusing task will be an important feature of an electro-optical imaging system that is to be used under high workload conditions.

It will be shown here, how the focusing difficulty may be quantified when the need for focusing arises solely from variations in object-range. (Effects due to variations in the optical system itself, for example as a function of temperature, are excluded.)

A fixed optical system is in theoretically perfect focus for only one value of the object range. However, an electro-optical imaging system will suffer resolution limitations other than defocus blur. A finite spread of object ranges will therefore exist for which the defocus blur will be an acceptably small fraction of the other resolution limitations. This spread of object ranges - over which the system can be considered to be "in-focus" - is called the depth-of-field of the optical system, for the particular range at which it is exactly focused.

The operational requirements for the imaging system will specify a minimum range (r) and a maximum range (R) between which the system is required to be "in-focus." The range ratio (σ) is defined as:

$$\sigma = R/r$$

If the depth of field is large enough, it may be possible to meet these range requirements with a single fixed-focus state of the optical system. If not, several fixed-focus positions will be needed. The focusing difficulty of the system may be quantified in terms of the minimum number (M) of such fixed-focus positions required to keep an object "in-focus" over the required spread of operating ranges. (This measure (M) quantifies the focusing difficulty even when the focusing control is continuous.)

Since M = 1 corresponds to zero focus-difficulty (a single fixed-focus condition sufficing for all required ranges), and since the degree of difficulty seems, intuitively, to be a logarithmic, rather than a linear, function of M it is proposed to define the Focus-Difficulty (F_D) as

$$F_D = \log_2 M$$

The optical system will be assumed to be diffraction limited, operating at wavelength λ. The optical system can be further defined in one of two ways:

(1) Directly - that is, the objective diameter (A) is specified;

(2) Indirectly - it is specified that the diffraction cut-off spatial frequency, at the object, must never be less than W cycles per meter over the specified spread of operating ranges.

The first case is appropriate when estimating the focusing difficulty of an existing imaging system, or of a new system for which resolution is specified in angular (not spatial) units. The second case arises when the focusing difficulty associated with the performance of a specific target detection or recognition task is to be investigated. That is, when the resolution is specified in spatial units at the object, irrespective of range.

This second method of specifying resolution will generally correspond to the real operational requirements of the system. Practical limitations on the diameter of the objective may, however, limit the maximum range (R) for which the spatial resolution specification can be met.

An important aspect of the spatial specification of resolution is that it allows the focusing difficulty to be minimized by the use of two, or more, different objectives to cover the required spread of operating ranges.

The principal results obtained in this section are summarized below (giving the focusing difficulty F_D in each case):

(The functions $F_N(\sigma)$, $P_{n,\,n+1}(F_N(\sigma))$ are defined and are also given in graphical form.

(1) Angular Specification of Resolution:

$$F_D = \log_2\left(\frac{\sigma - 1}{\sigma} \quad \frac{\delta}{r}\right)$$

(2) Spatial Specification of Resolution (single objective):

$$F_D = \log_2((\sigma - 1)\,\frac{R}{S})$$

$$A = A_o R/S \text{ (diameter of the objective)}$$

(3) Spatial Specification of Resolution (lens turret with N objectives):

$$F_{DN} = \log_2\,(N\,\frac{R}{S}\,F_N(\sigma))$$

$$\left.\begin{aligned} A_1 &= A_o R/S \\ A_2 &= A_1/P_{12} \\ &\text{----------} \\ A_N &= A_{N-1}/P_{N,N+1} \end{aligned}\right\} \text{ diameters of the N objective}$$

(A value of $F_D < 0$ is to be interpreted as zero focusing difficulty.)

The associated parameters are:

x = criterion for acceptable in-focus (typically, $x = 1$)	operational requirements
r = minimum operating range	
R = maximum operating range	
$\sigma = R/r$ (range - ratio)	
A = diameter of objective lens	resolution requirements
W = cut-off frequency of diffraction limited MTF in cycles/meter at the object	
$\delta = A^2/2x\lambda$ (semihyperfocal distance)	derived parameters
$S = 2x/\lambda W^2$ (critical focus distance)	
$A_o = 2x/W$ (critical aperture)	

In the following analysis the specification of an acceptable defocus blur is considered first. Then, the minimum number (M) of fixed focus steps required to meet both types of resolution specifications are derived as a function of the range ratio σ, and of either the minimum range (r) or of the maximum range (R).

4.6.2 Specification of an Acceptable Defocus Blur

The MTF of defocused diffraction limited optics is shown in Figure 4.2, taken from Reference 4.1. The parameter Δ is a measure of the focus error, and is given by

$$\Delta = 2 \sin^2 \left(\frac{A}{2f}\right) \frac{\Delta i}{\lambda}$$

$$\approx \frac{A^2}{2f^2} \frac{\Delta i}{\lambda} \qquad (4.1)$$

where A is the diameter, f the focal length of the objective, and Δi is the axial focus error at the image plane.

The diameter, d, of the geometrical defocus blur circle is given by

$$d = \frac{\Delta i \, A}{f}$$

The size of the (sharply focused) diffraction image of a point is proportional to $f\lambda/A$ (the Airy Disc diameter is $2.22\ f\lambda/A$).

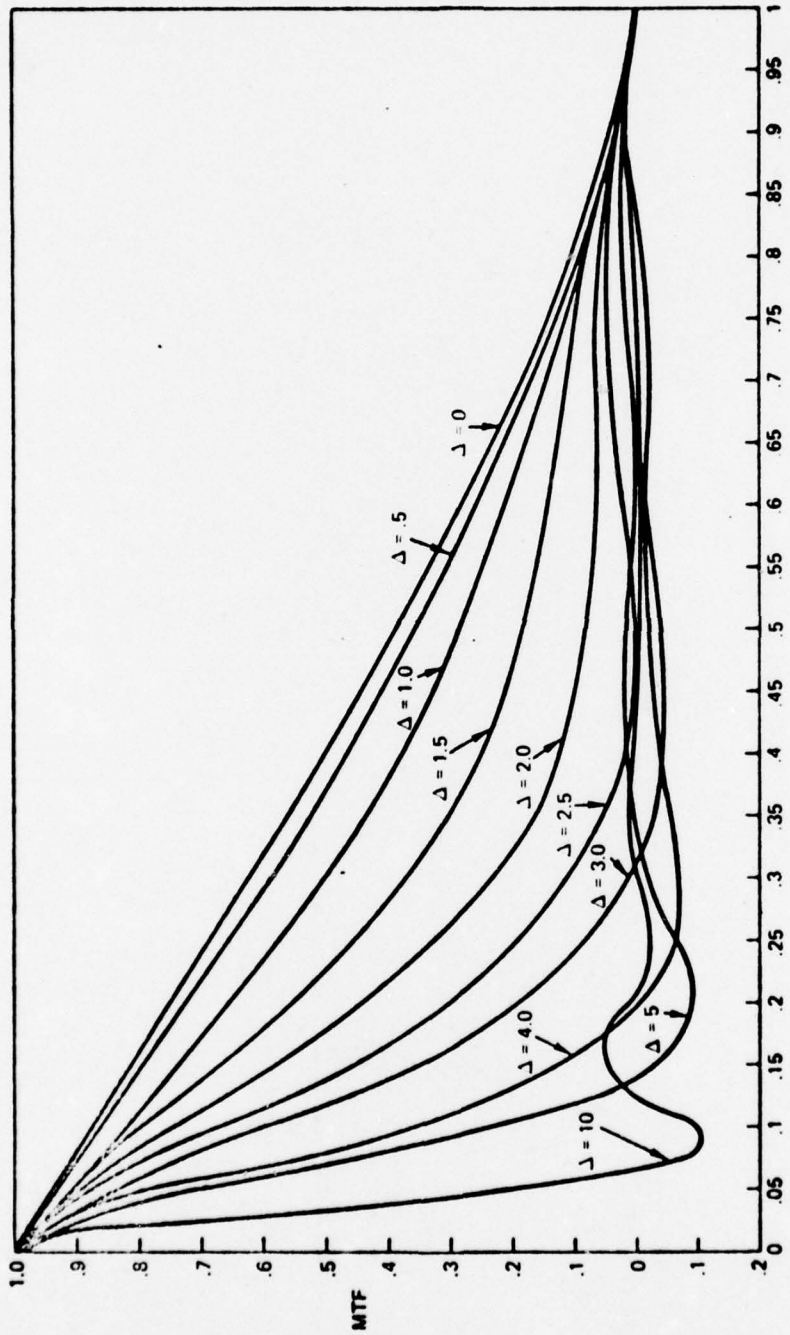


Figure 4.2 MRF's FOR DEFOCUSED DIFFRACTION-LIMITED OPTICS
ADAPTED FROM REFERENCE 4.1

It is convenient to express the degree of defocus as a ratio (x) of the defocus blur to the diffraction blur (Figure 4.3). That is,

$$x = \frac{d}{f\lambda/A} = \frac{\Delta i\ A/f}{f\lambda/A}$$

$$= \frac{\Delta i\ A^2}{f^2\lambda} \tag{4.2}$$

From equations (4.1) and (4.2):

$$x = 2\Delta \tag{4.3}$$

The MTF of defocused diffraction limited optics is also discussed, in Reference 4.2, in terms of the optical path difference (OPD), where:

$$OPD = \frac{1}{2}\{\sin^2 (A/2f)\}\Delta i$$

$$= \frac{A^2}{8f^2}\Delta i$$

Substituting for Δi from equation (4.1)

$$OPD = \frac{\lambda\Delta}{4} \tag{4.4}$$

It can be seen from Figure 4.2 that $\Delta = 1$ is a suitable infocus criterion for most applications:*

(1) the resultant MTF degradation is not negligible (which would lead to an unrealistically small depth-of-field), yet

(2) it does not represent a drastic reduction in MTF.

From equations (4.3) and (4.4) it follows that this ($\Delta = 1$) in-focus criterion also implies:

$$OPD = \lambda/4$$

and

$$x = 2$$

The overall MTF of an electro-optical imaging system is made up of several components (for example: detector, electronics, optics).

*For targets whose size and contrast are such that they are only just detectable in a sharply focused ($\Delta = 0$) image, the $\Delta = 1$ criterion (which represents a loss of about 25 percent in contrast) would not be acceptable. If it is important that such targets not be missed, a Δ criterion of 0.5 or even of 0.25 should be used.

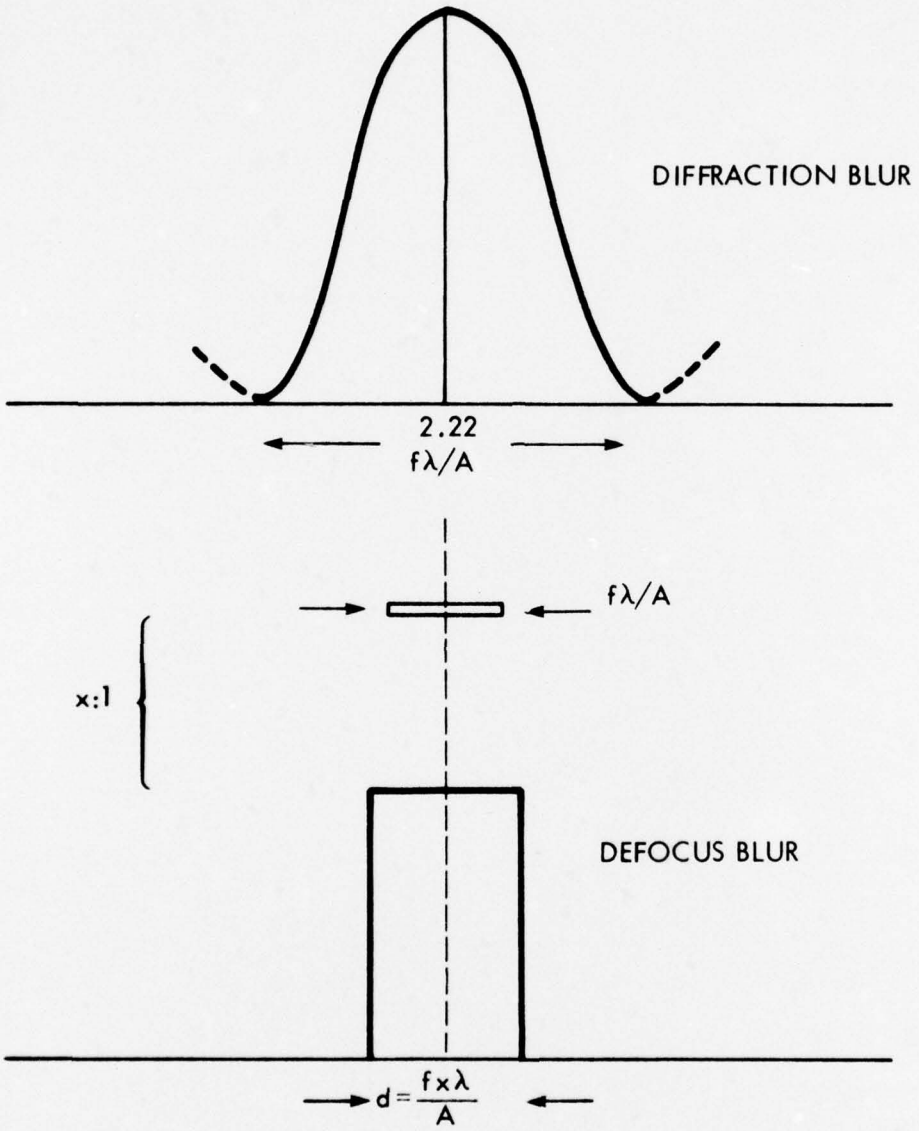


Figure 4.3 DEFINITION OF x

However, because MTF-degradation caused by defocus (over the range of practical interest of $0 < \Delta < 2$) extends fairly uniformly over the full range of spatial frequencies the practical effect (or significance) of any such degree of defocus would be largely independent of the other MTF's of the system (for example, the detector MTF).

Accordingly an "in-focus" criterion, based upon a particular Δ (or x) value, can be considered as equally valid over the practical range of the other component MTF's.

4.6.3 Depth-of-Field Analysis

4.6.3.1 Angular Specification of Resolution

Let the angular size of the acceptable geometric defocus blur cycle be β. Then by definition of x (Figure 4.4),

$$\beta = \frac{x\lambda}{A} \tag{4.4}$$

If an optical system is focused at the hyperfocal distance (D_{HF}) it is acceptably in-focus from $D_{HF}/2$ out to infinity. The hyperfocal distance is given by:

$$D_{HF} = A/\beta$$

or, from (4),

$$D_{HF} = A2/x\lambda \tag{4.5}$$

A finite hyperfocal distance, D, can also be defined. It is the focusing distance for which the angular size of the geometrical defocus blur is acceptable (that is less than β) for all points from a minimum range r to a maximum range R. From Figure 4.4, at the maximum range (R),

$$\beta = \frac{A(R - D)}{DR}$$

i.e.,

$$D = \frac{AR}{A + R\beta}$$

$$= \frac{R}{1 + R/D_{HF}} \tag{4.6}$$

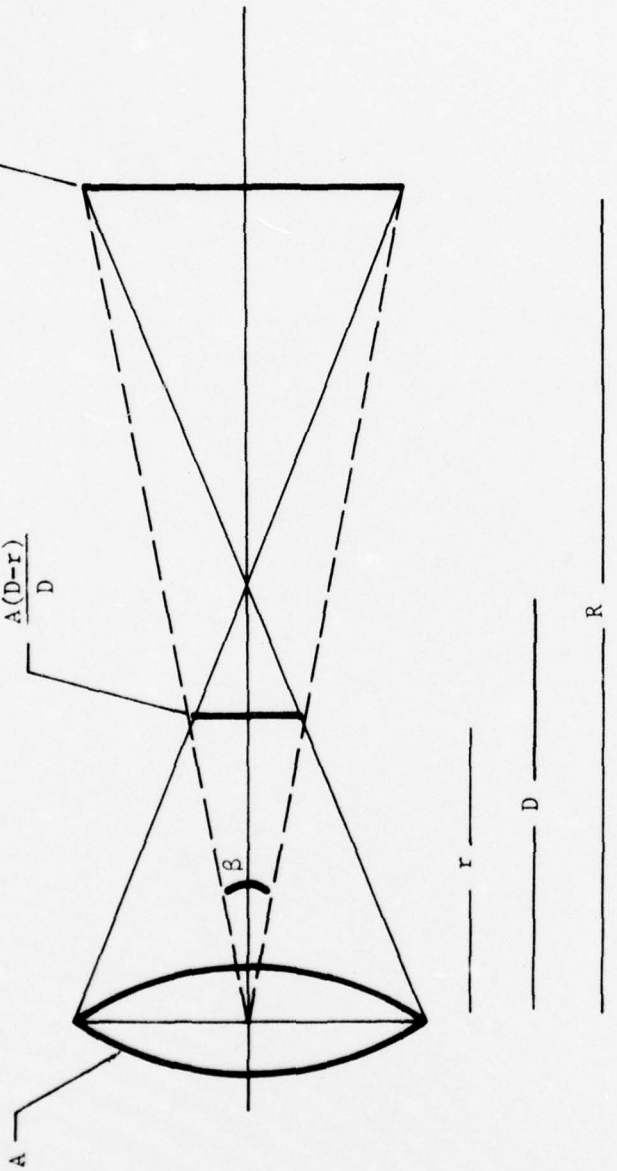


Figure 4.4 FINITE HYPERFOCAL DISTANCE

Likewise, at the minimum range (r),

$$\beta = \frac{A(D - r)}{Dr}$$

i.e.,

$$r = \frac{AD}{D\beta + A}$$

$$= \frac{D}{1 + D/D_{HF}} \qquad (4.7)$$

Substituting for D from equation (4.6) into (4.7)

$$r = \frac{R}{1 + 2R/D_{HF}}$$

It is convenient to introduce the semihyperfocal distance (δ),

$$\delta = D_{HF/2} = \frac{A}{2\beta} = \frac{A^2}{2x\lambda} = \qquad (4.8)$$

(Note that δ is a constant [for a given wavelength] for a given <u>angular</u> specification of acceptable defocus blur.)

Then from equations (4.6) and (4.8);

$$D = \frac{R}{1 + R/2\delta} \qquad (4.9)$$

and from equations (4.7) and (4.8);

$$r = \frac{R}{1 + R/\delta}$$

Since

$$R/R = \delta$$

it follows from the last equation that

$$\delta = 1 + R/\delta \qquad (4.10)$$

This equation gives the range ratio (σ) that can be achieved with a single fixed focus condition of the optics for a given maximum range (R). The in-focus range ratio (σ) that can be achieved for a given minimum range (r) is given by

$$\sigma = 1 + \sigma r/\delta$$

i.e.,

$$\sigma = \frac{\delta}{\delta - r} \tag{4.11}$$

Equations (4.10) and (4.11) show the practical significance to the semihyperfocal distance. For ranges less than the semihyperfocal distance, in-focus range-ratios are limited and range-focusing may be necessary.

The range-ratio (σ) covered by a single fixed-focus condition has been derived. The range ratio that can be achieved with multiple fixed focus conditions will now be determined.

Let R_1 be the maximum range at which it is desired to be "in-focus." By focusing at a range of

$$D_1 = \frac{R_1}{1 + R_2/2\delta}$$

the system is in-focus from ranges R_1 down to r_1 where

$$r_1 = \frac{R_1}{1 + R_1/\delta}$$

Consider now, a second fixed-focus position in which the new maximum range (R_2) is chosen to be

$$R_2 = r_1$$

so that the second "in-focus" region takes over at the end of the first region.

The second fixed focus distance should be:

$$D_2 = \frac{R_2}{1 + R_2/2\delta}$$

and the system will then be in-focus from R_2 to r_2, where

$$r_2 = \frac{R_2}{1 + R_2/S}$$

As shown in Figure 4.5

$$r_2 = \frac{R_2}{1 + 2R_1/\delta}$$

$$D_2 = \frac{R_1}{1 + 3R_1/2\delta}$$

For M such fixed-focus steps

$$r = \frac{R}{1 + MR/\delta}$$

$$D = \frac{R}{1 + \frac{2M + 1}{2}\frac{R}{\delta}} \qquad (4.12)$$

In other words, the range ratio (σ) that can be achieved with M fixed-focus steps is:

$$\sigma = 1 + \frac{MR}{\delta}$$

or

$$\sigma = \frac{1 + M\sigma_r}{\delta}$$

i.e.,

$$M = \frac{(\sigma - 1)}{\sigma}\,\delta/r \qquad (4.13)$$

The focusing difficulty, F_D, is then given by

$$F_D = \log_2\left(\frac{\sigma - 1}{\sigma}\,\frac{\delta}{r}\right) \qquad (4.14)$$

This function is plotted in Figure 4.6.

In using Figure 4.6, the semihyperfocal distance (δ) should be calculated first. Then the focusing difficulty, for any minimum range relative to δ, can be read-off directly for the desired range ratio (σ).

Figure 4.6 illustrates that

(1) for minimum-ranges greater than the semihyperfocal distance, the focusing difficulty is zero, and

(2) for minimum-ranges less than semihyperfocal distance from the focusing difficulty increases as the range decreases.

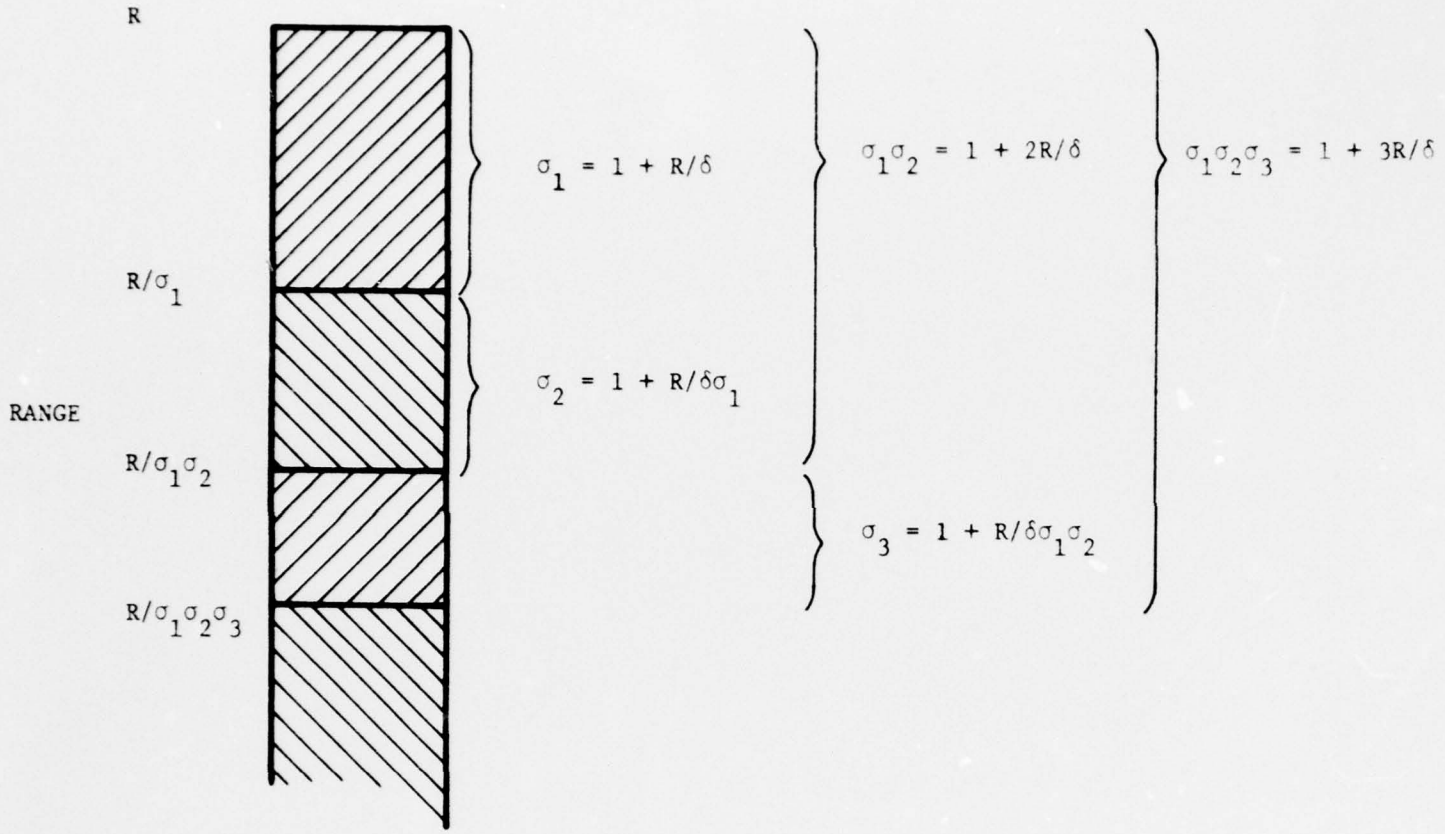


Figure 4.5 SERIES OF FIXED-FOCUS STEPS

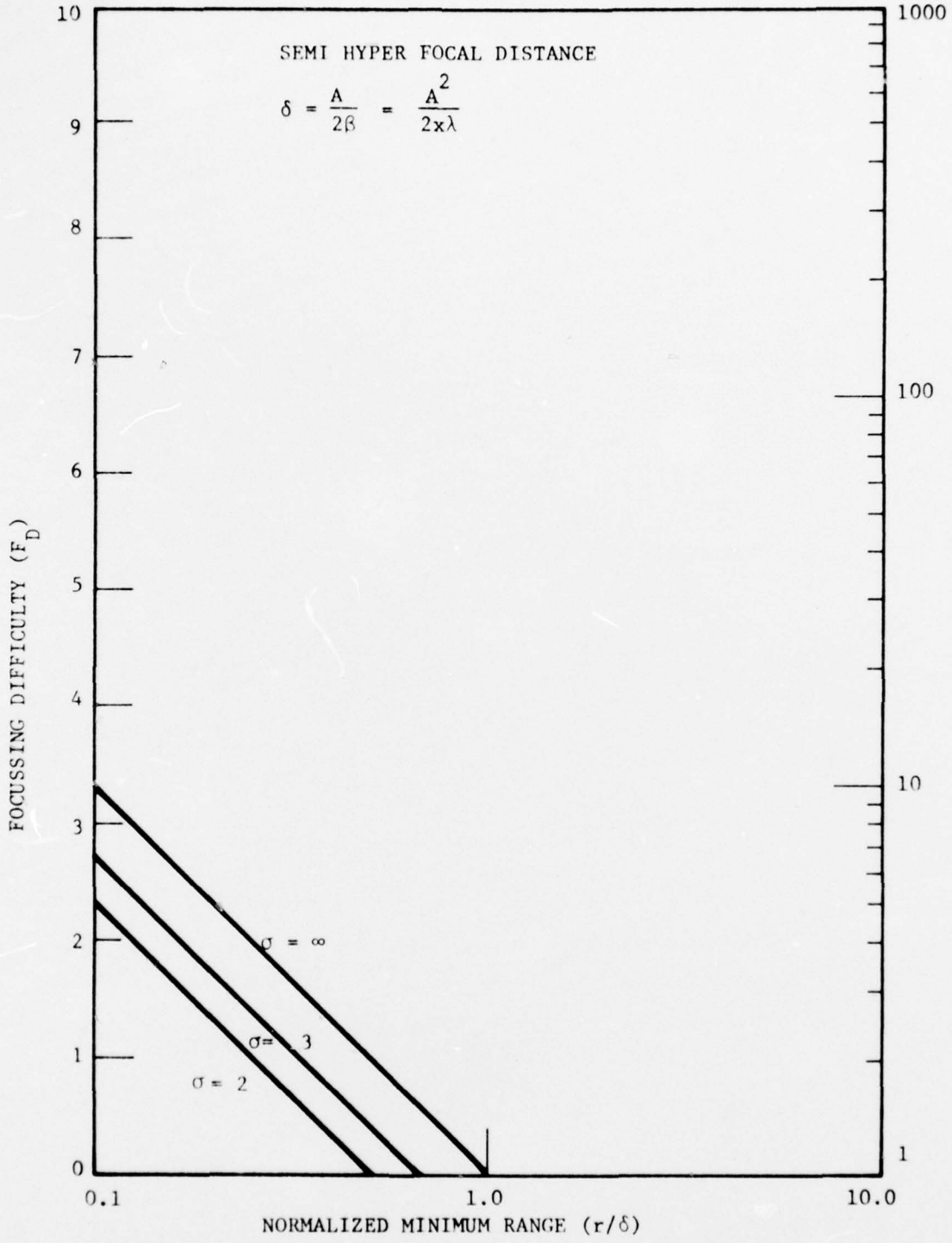


Figure 4.6 FOCUSING DIFFICULTY: ANGULAR RESOLUTION SPEC.

77-5-1

Thus the semihyperfocal distance can be regarded as the boundary-range between regions of "easy focusing" and "difficult focusing."

The value of the semihyperfocal distance (δ) is plotted in Figure 4.7 as a function of objective diameter (A) for $\lambda = 4\mu m$ and $\tau = 10\mu m$.

4.6.3.1.1 Upper Bound Estimate of Focusing Difficulty

An upper-bound estimate $(F_D)_2$ of the focusing difficulty can be derived from equation (4.14) by setting $\sigma = \infty$:

$$(F_D)_L = \log_2 \left(\frac{\delta}{r}\right)$$

where r is the minimum range at which the system will be used.

A minimum value for the range at which the system is likely to be used can be estimated as the range at which the system IFOV can resolve the smallest detail of interest on the target.

(The implicit assumption is that, at shorter ranges than this, a wider field-of-view objective would normally be used. However, if the system is used at shorter ranges, without changing the objective, the operator may then keep trying to focus the system, even though the defocus blur circle is less than the smallest detail of practical interest. That is, he may exert effort in focusing for aesthetic, rather than practical, reasons. In this event the size of the smallest detail of "interest" must be appropriately reduced.)

Let x be the smallest required spatial dimension of the IFOV at the target. Then

$$r = x/\alpha$$

where r is the minimum range and α is the angular IFOV of the system.

The angular IFOV (α) is related to the diffraction limit of the system:

$$\alpha = \frac{K\lambda}{A}$$

where K is a design choice, usually satisfying

$$1 \leq K \leq 2$$

From equation (4.8)

$$\sigma = \frac{A^2}{2x\lambda}$$

$$\therefore (F_D)_L = \log_2\left(1 + \frac{A^2}{2x\lambda}\frac{K\lambda}{AX_o}\right) = \log_2\left(1 + \frac{K}{2x}\frac{A}{X_o}\right)$$

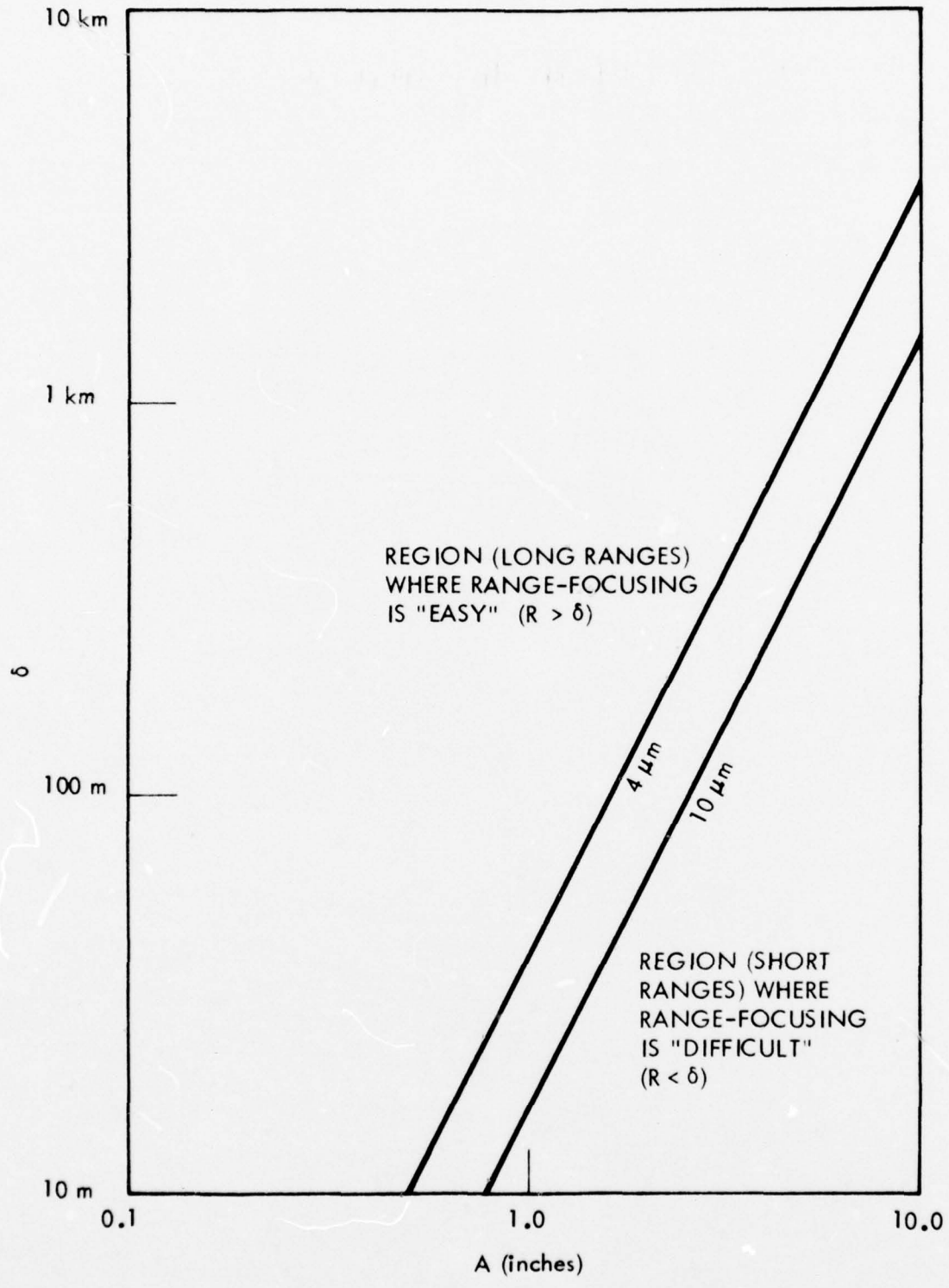


Figure 4.7 SEMI HYPERFOCAL DISTANCE
(ASSUMES x = 1, THAT IS Δ=0.5)

77-5-1

(In this equation, K and x are not totally independent. For relatively large values of K (e.g., $\sim$2) the overall system -MTF is more detector-limited than diffraction-limited, and consequently a (relatively) larger value of x would be considered as an acceptable in-focus criterion.)

Assuming typical values for K and x:

$$K = 2$$

$$x = 2$$

it follows that

$$(F_D)_L = \log_2 \left(1 + A/(2X_o)\right).$$

This equation gives an upper-bound estimate of the focusing difficulty as a function of the ratio of the objective diameter (A) to the finest spatial resolution required (X_o).

Typically, FLIR apertures are often limited to 8 inches or less by practical system constraints. For most targets a minimum spatial resolution of 4 inches would be adequate. In this case the upper-bound focusing difficulty is therefore unity (2 fixed focus steps).

4.6.3.2 Spatial Specification of Resolution

In this case, the diffraction MTF cut-off spatial frequency at the object is required to be W cycles/meter, at the maximum range (R)

i.e.,

$$W = \frac{A}{R\lambda} \quad c/m$$

i.e.,

$$A = \lambda RW \tag{4.15}$$

Then, by definition of x,

$$\beta = \frac{x\lambda}{A} = \frac{x}{WR} \tag{4.16}$$

Then δ is given by

$$\delta = \frac{A}{2} = \frac{\lambda R^2 W^2}{2x} = \frac{R^2}{S} \tag{4.17}$$

where S (defined here as the critical focus distance) is given by

$$S = \frac{2x}{\lambda W^2} \tag{4.18}$$

S is a parameter similar to δ (the semihyperfocal distance) in that it is a constant (for a given wavelength) for a given spatial specification of acceptable defocus blur. (δ is a constant for a given angular specification of defocus blur.)

Substituting for δ from equation (4.18) into equations (4.12) and (4.13)

$$D = \frac{R}{1 + \frac{2M + 1}{2} \frac{S}{R}} \tag{4.19}$$

$$\sigma = 1 + M\ S/R \tag{4.20}$$

These equations show how an "in-focus" range ratio σ can be achieved, for a maximum range R, when the resolution is specified spatially.

The spatial resolution is (just) met at the maximum range R and is exceeded at all other ranges. The angular size of the acceptable defocus blur is constant for all ranges between the minimum (r) and the maximum (R). Subjectively, therefore, the system will appear focused as an object moves from the maximum (R) to the minimum range (r).

The corresponding focusing difficulty is given by

$$F_D = \log_2 M = \log_2 \left\{(\sigma - 1) \frac{R}{S}\right\} \tag{4.21}$$

and is plotted in Figure 4.8.

Also shown in Figure 4.8 is the corresponding objective diameter (A), given by

$$A = \lambda RW$$

$$= \lambda SW \frac{R}{S}$$

$$= \frac{2x}{W} \frac{R}{S}$$

$$= A_o \frac{R}{S}$$

i.e.,

$$A/A_o = R/S \tag{4.22}$$

where A_o (defined here as the critical aperture diameter) is given by

$$A_o = \frac{2x}{W} \tag{4.23}$$

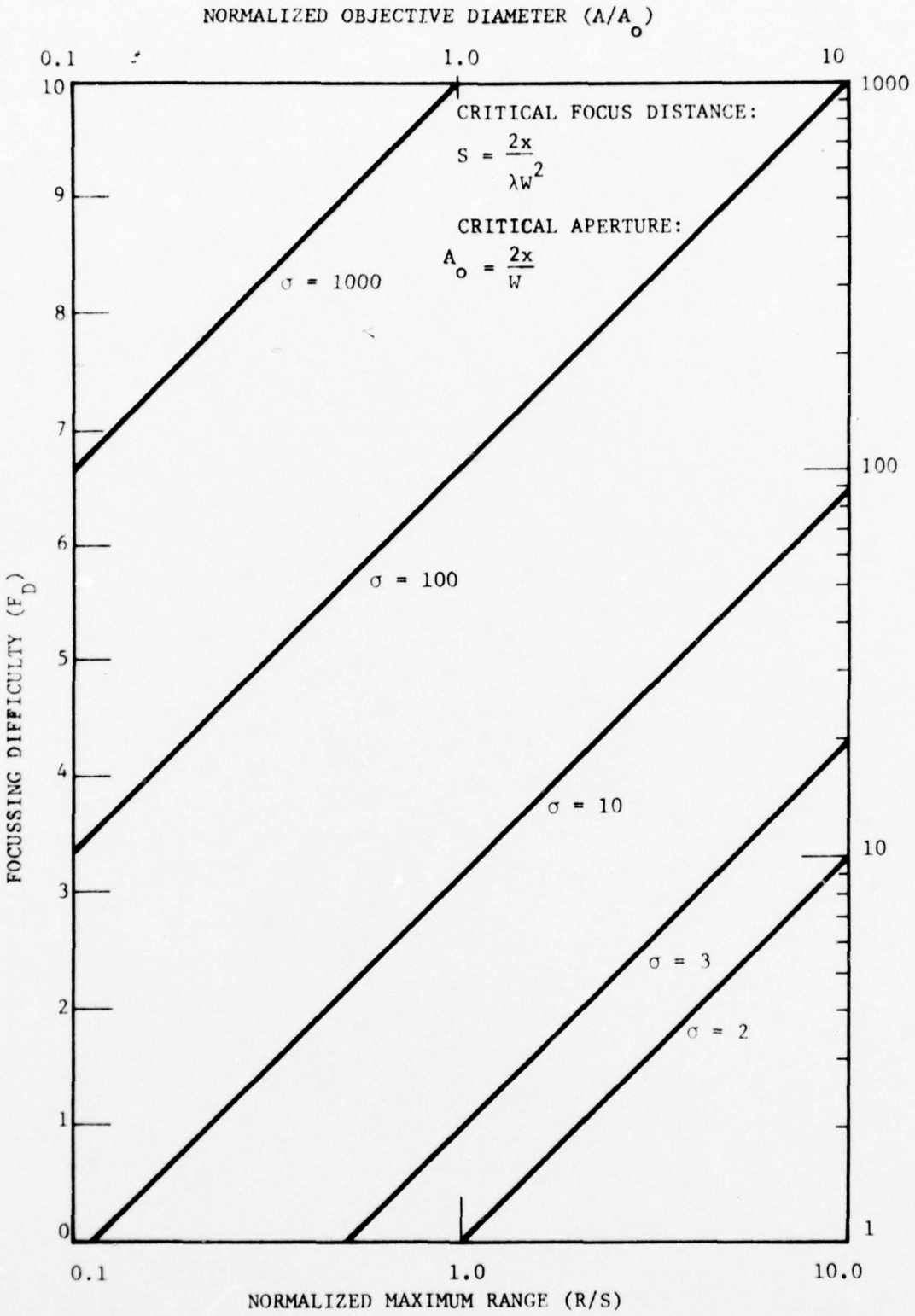


Figure 4.8 FOCUSING DIFFICULTY: SPATIAL RESOLUTION SPEC (SINGLE OBJECTIVE)

In using Figure 4.8, the critical focus distance S should be calculated first. Then the focusing difficulty can be read off directly, for any maximum range (R) relative to S, for any desired range ratio σ.

Figure 4.8 shows the practical significance of the critical focus distance (S). For ranges less than the critical focus distance, range-focusing (with a spatial specification of resolution) is relatively easy. For ranges greater than the critical focus distance, range-focusing is relatively difficult. It may be noted that this is the exact opposite of the case for range-focusing with an angular specification of resolution (Figure 4.6).

The value of the critical focus distance (S) is plotted in Figure 4.9 as a function of the spatial resolution (W c/m), for λ = 4μm and λ = 10μm. (The values of W corresponding to detection, orientation, recognition, identification of a target having a minimum dimension of 2m are also marked.)

When the resolution requirements are spatial, the focusing difficulty can be reduced through the use of a series of objectives of progressively smaller apertures (A1, A2, A3 . . .), instead of just one objective. The smaller objectives are used at the shorter ranges. Each objective will be assumed to have the same F #, so that the diffraction blur, as seen by the electro-optical detector, will appear the same in each case.

Let the first objective be used at the maximum range R to provide a range ratio σ_1:

$$\sigma_1 = 1 + MS/R$$

$$= 1 + \rho$$

where

$$\rho = MS/R$$

The second objective will be used, from a maximum range R/σ_1, to provide a range ratio σ_2:

$$\sigma_2 = 1 + \frac{MS}{R/\sigma_1}$$

$$= 1 + \sigma_1 \, MS/R = 1 + \sigma_1\rho$$

$$= 1 + G_1\rho$$

where

$$G_1 = \sigma_1$$

(It is assumed for simplicity, that the same number (M) of fixed-focus steps is used for each objective.)

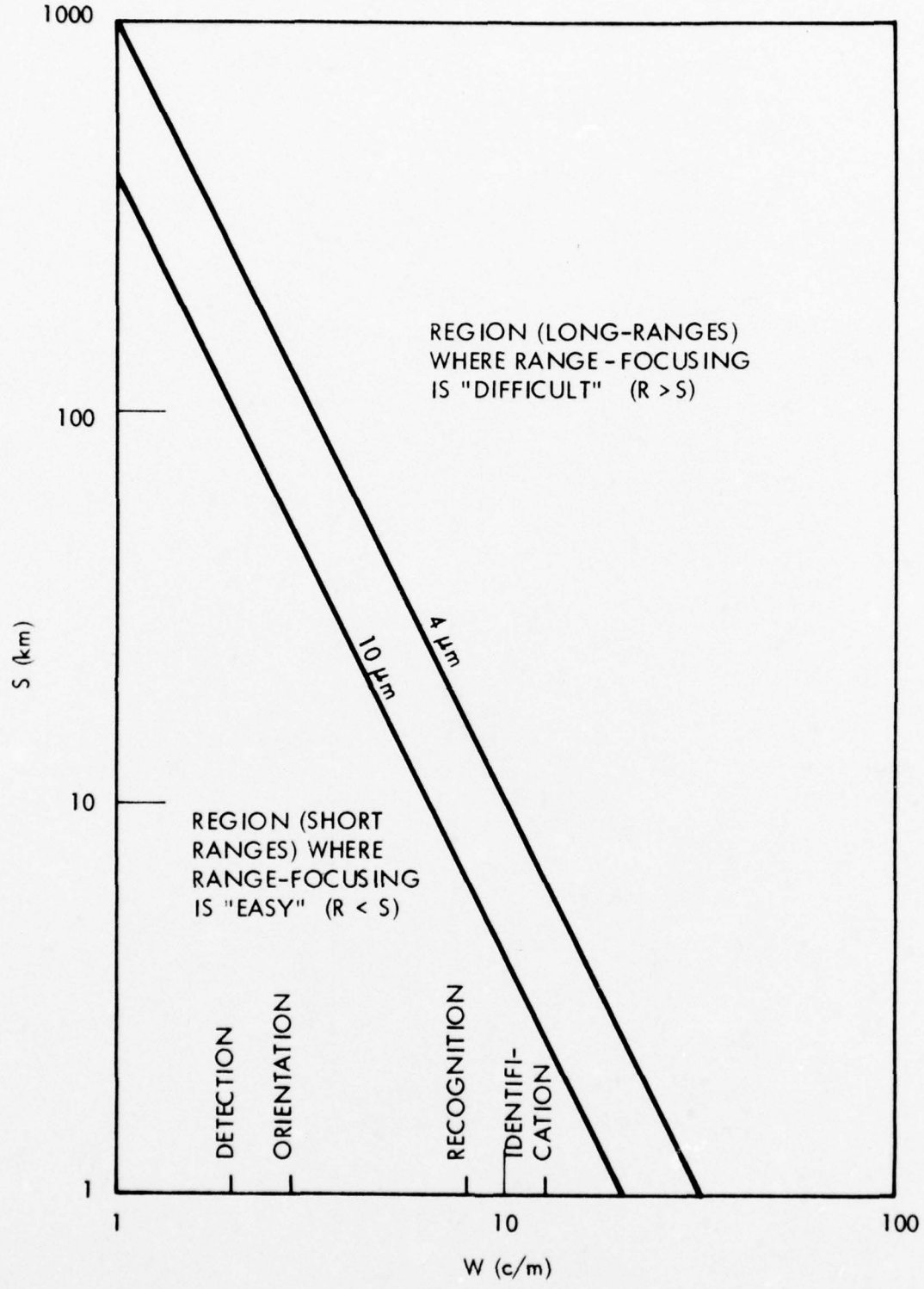


Figure 4.9 CRITICAL FOCUS DISTANCE(S)
(ASSUMES x = 1)

77-5-1

The overall range ratio (G_2) for the two steps is

$$G_2 = \sigma_1\sigma_2 = G_1(1 + G_1\rho)$$

In general, after (n) steps,

$$f_n = G_{n-1}(1 + G_{n-1}\rho)$$

$$\sigma_n = 1 + G_{n-1}\rho = P_{n,n+1}$$

where $P_{n,n+1}$ is the power ratio of the nth and n+1th objectives. The function

$$\sigma = G_N(\rho)$$

where σ is the overall range ratio, can be expressed as

$$\rho = F_N(\sigma) \tag{4.24}$$

Then, $P_{N,N=1}$ can be considered as a function of ρ.

The functions $P_{N,N+1}(\rho)$, $F_N(\sigma)$ are plotted in Figure 4.10 for values of N from 1 to 6. The corresponding focus-difficulty functions can be derived from equation (4.24):

$$\rho = MS/R = F_N(\sigma)$$

i.e.,

$$M = F_N(\sigma)\ R/S$$

where M is the number of fixed focus steps per objective. Since there are N objectives the total number of steps in NM. Thus,

$$F_{DN} = \log_2 (NM)$$

$$= \log_2 (NR/S\ F_N(\sigma)) \tag{4.25}$$

(This derivation of the focus-difficulty assumes that an objective change is equally "difficult" as a change to the next fixed-focus position with the same objective. This is probably true in terms of the operator intervention required to command the change. However, an objective change may be more disturbing to the operator. In this sense the focusing "difficulty" may be somewhat underestimated by equation (4.25).)

The Focus Difficulty function (F_{DN}) is plotted in Figures 4.8, 4.11 and 4.12 for the cases of one, two and three objectives.

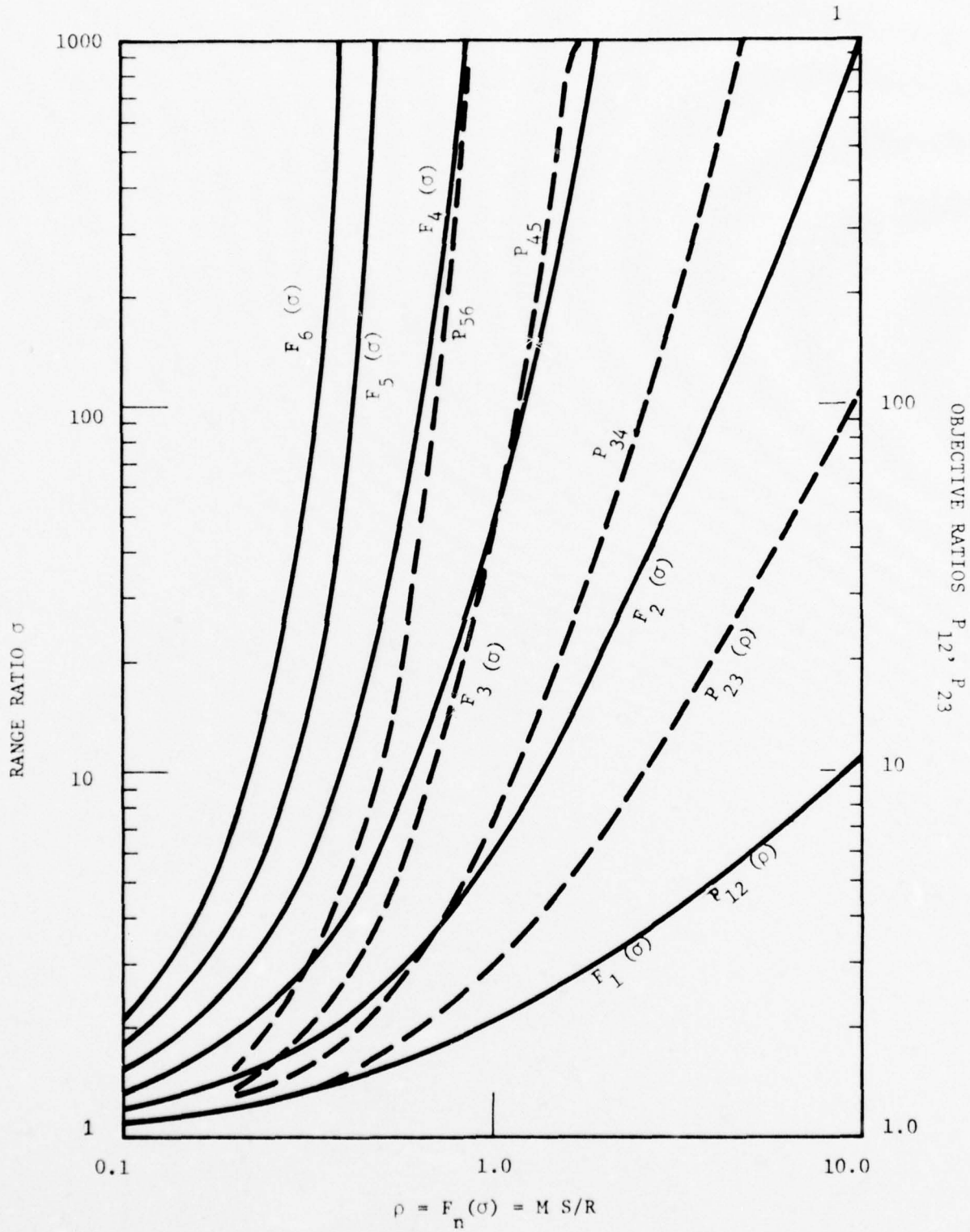


Figure 4.10 RANGE RATIOS FOR UP TO SIX OBJECTIVES

77-5-1

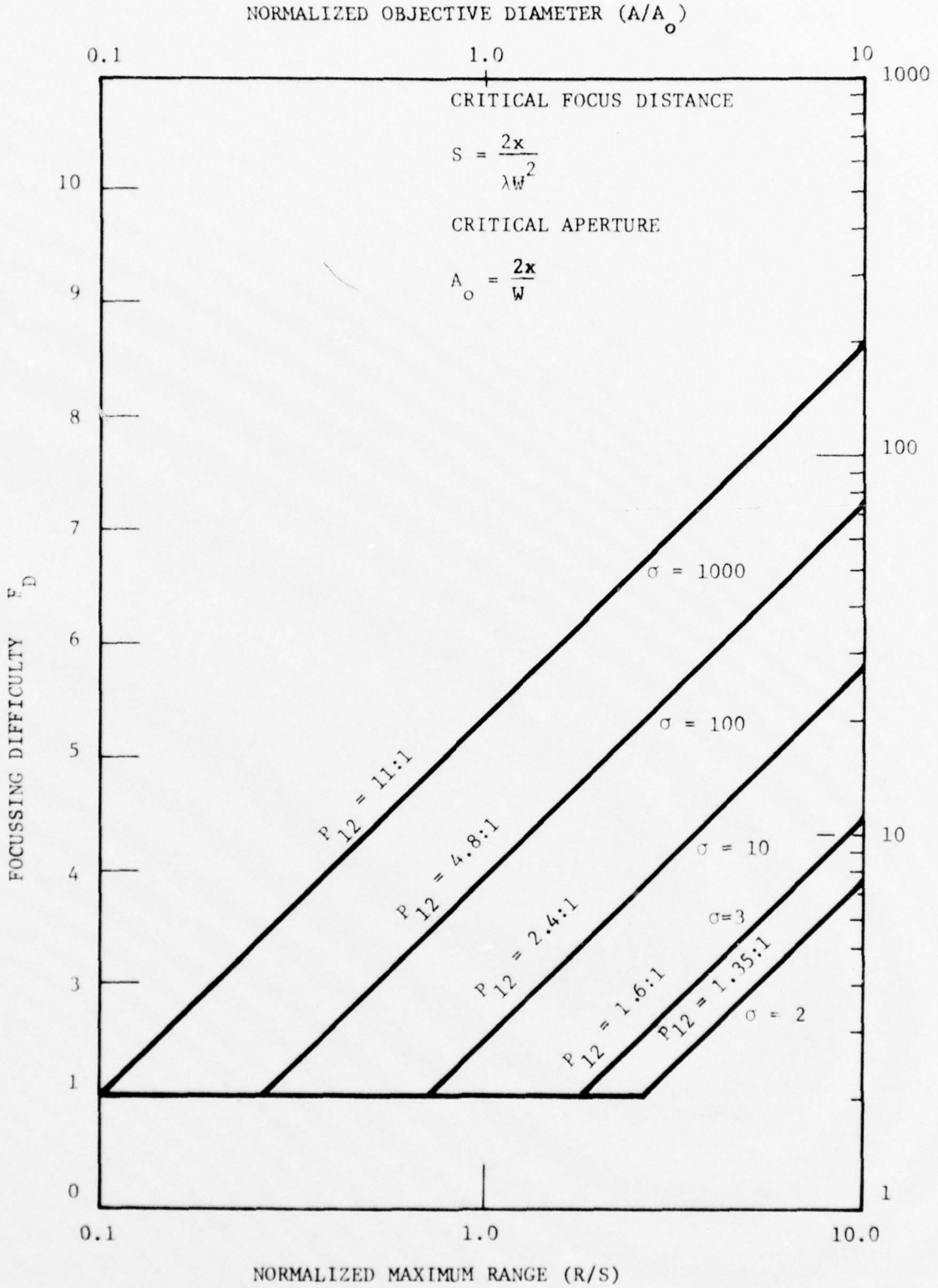


Figure 4.11 FOCUSING DIFFICULTY: SPATIAL RESOLUTION SPEC (TWO OBJECTIVES)

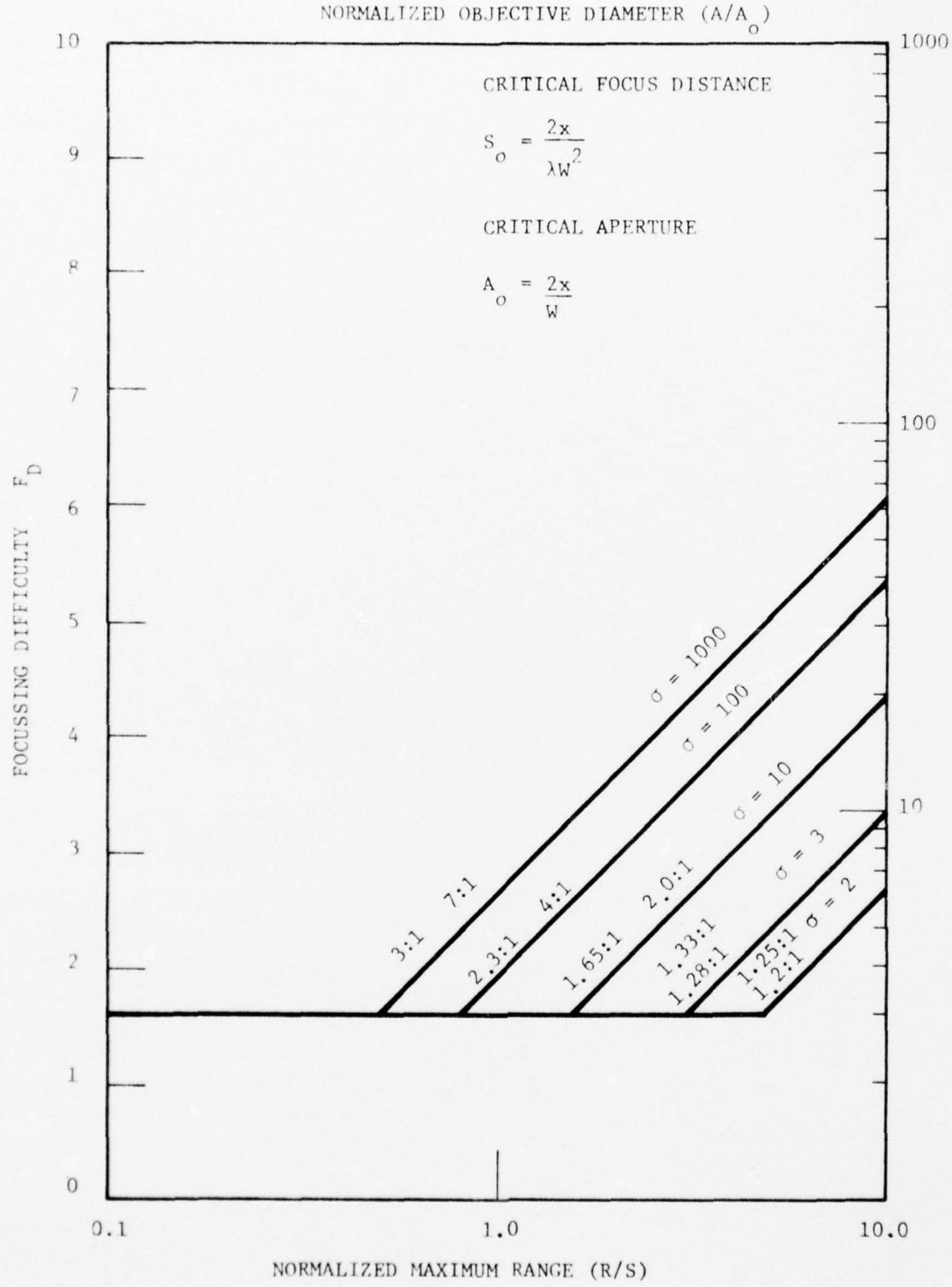


Figure 4.12 FOCUSING DIFFICULTY: SPATIAL RESOLUTION SPEC (THREE OBJECTIVES)

77-5-1

The focusing difficulty has been computed in terms of resolution at the object, in the form of the diffraction cut-off spatial frequency W cycles per meter. The advantage of this spatial resolution specification is that it can be related to specific target detection functions. For this purpose it is assumed that one TV-line corresponds to 2/W meters at the target, where W is the diffraction cut-off spatial frequency in cycles per meter. Table 4.1 gives the value of W needed for the detection, orientation, recognition, identification, of a target having a minimum dimension of 2 meters according to Johnson's criteria (Reference 4.3).

The focusing difficulty, F_D, in achieving these four levels of target discrimination are shown in Figure 4.13, for λ = 4μm and for λ = 10μm, assuming the use of up to three objectives.

To illustrate the use of this figure, suppose that it is desired to estimate the focusing difficulty associated with the function of "recognition," using a 4μm FLIR at 10 km. The range scale for this particular example is the one marked "R" (for recognition), and 4μm. (The eight range scales are shown at the top of the figure.)

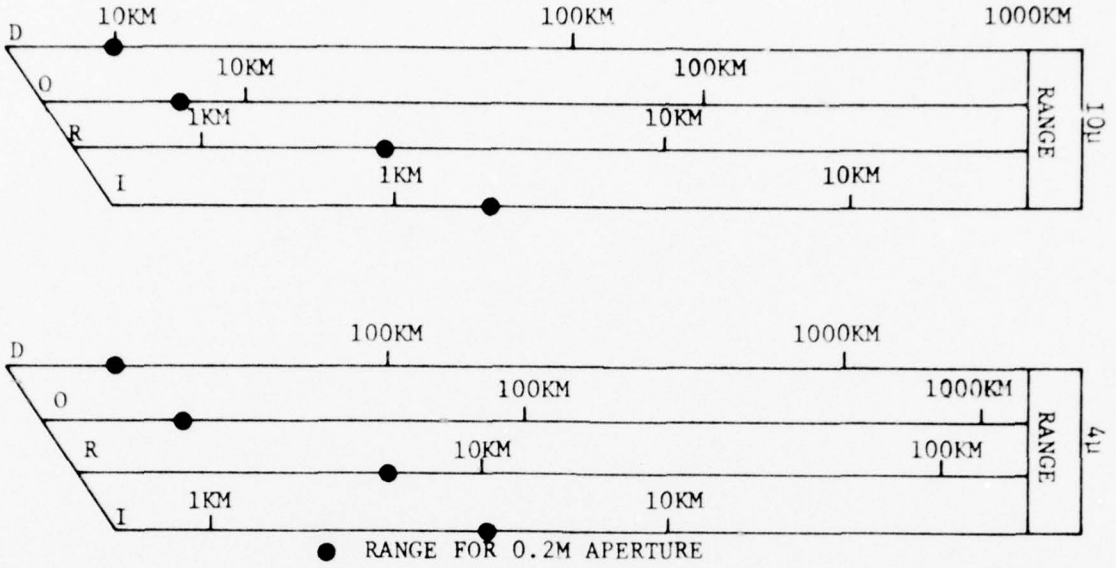
It can now be seen, from the figure, that the focusing difficulty for a maximum range of 10 km is 2, for a range-ratio of 1000, and 1, for a range-ratio of 10.

The range scales also give an indication of the corresponding objective diameter required to meet the specifications. This is done by means of a mark showing the range corresponding to a 0.2m diameter objective. If, as in the present example, a greater range (i.e., 10 km) is needed, then the objective diameter must be proportionally larger (than 0.2m).

The focusing difficulty assuming the use of up to six objectives is shown in Figure 4.14, and shows the trend as the number of objectives is increased. It can be seen that the use of multiple objectives substantially reduces the focusing difficulty with the biggest reductions obtained on conversion from a single to a dual, or triple, objective system.

TABLE 4.1
DETECTION CRITERIA FOR A 2m MINIMUM DIMENSION TARGET

Discrimination Level	TV Lines per min Dimension	W c/m	(A_o) Critical Aperture (M)	(S) Critical Focus Distance km	
				λ = 4 μm	λ = 10 μm
Detection	2	2	2	250	100
Orientation	2.8	2.8	1.43	127	51
Recognition	8	8	0.5	15.6	6.25
Identification	12.8	12.8	0.31	6.1	2.44



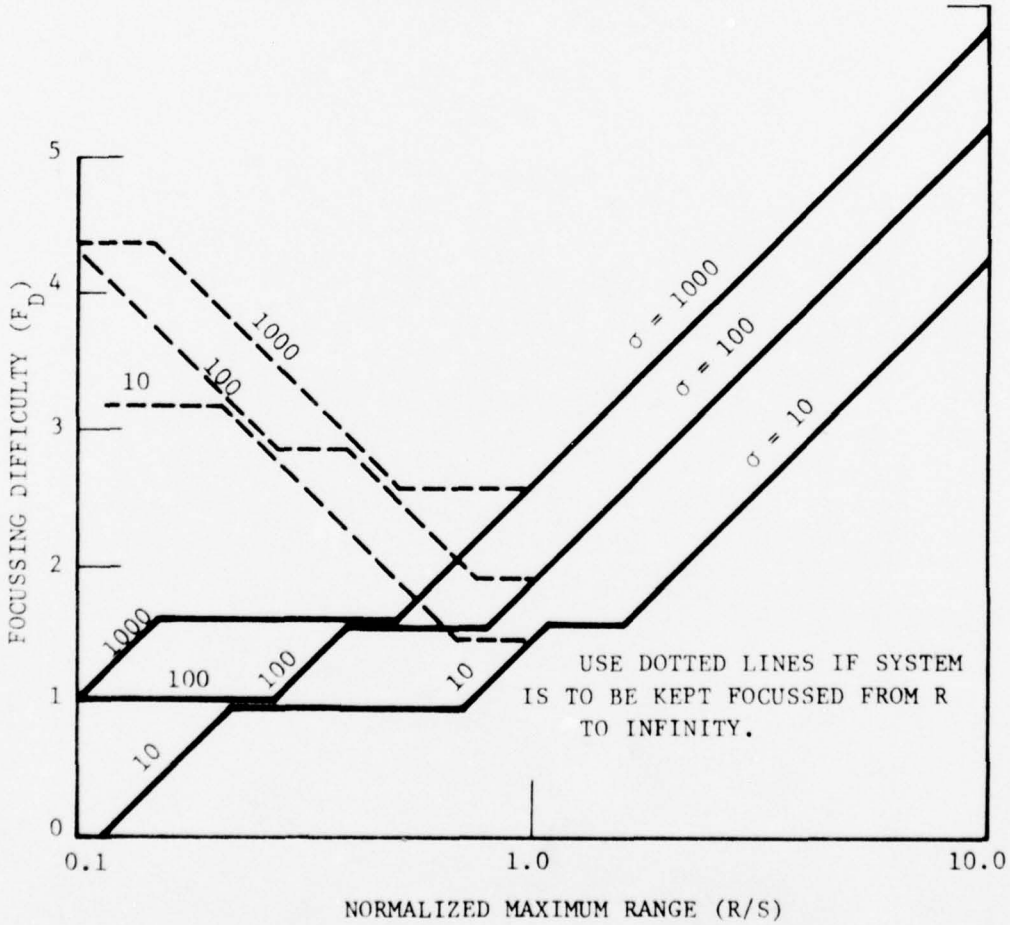


Figure 4.13 FOCUSING DIFFICULTY USING UP TO THREE OBJECTIVES (SPATIAL RESOLUTION SPEC)

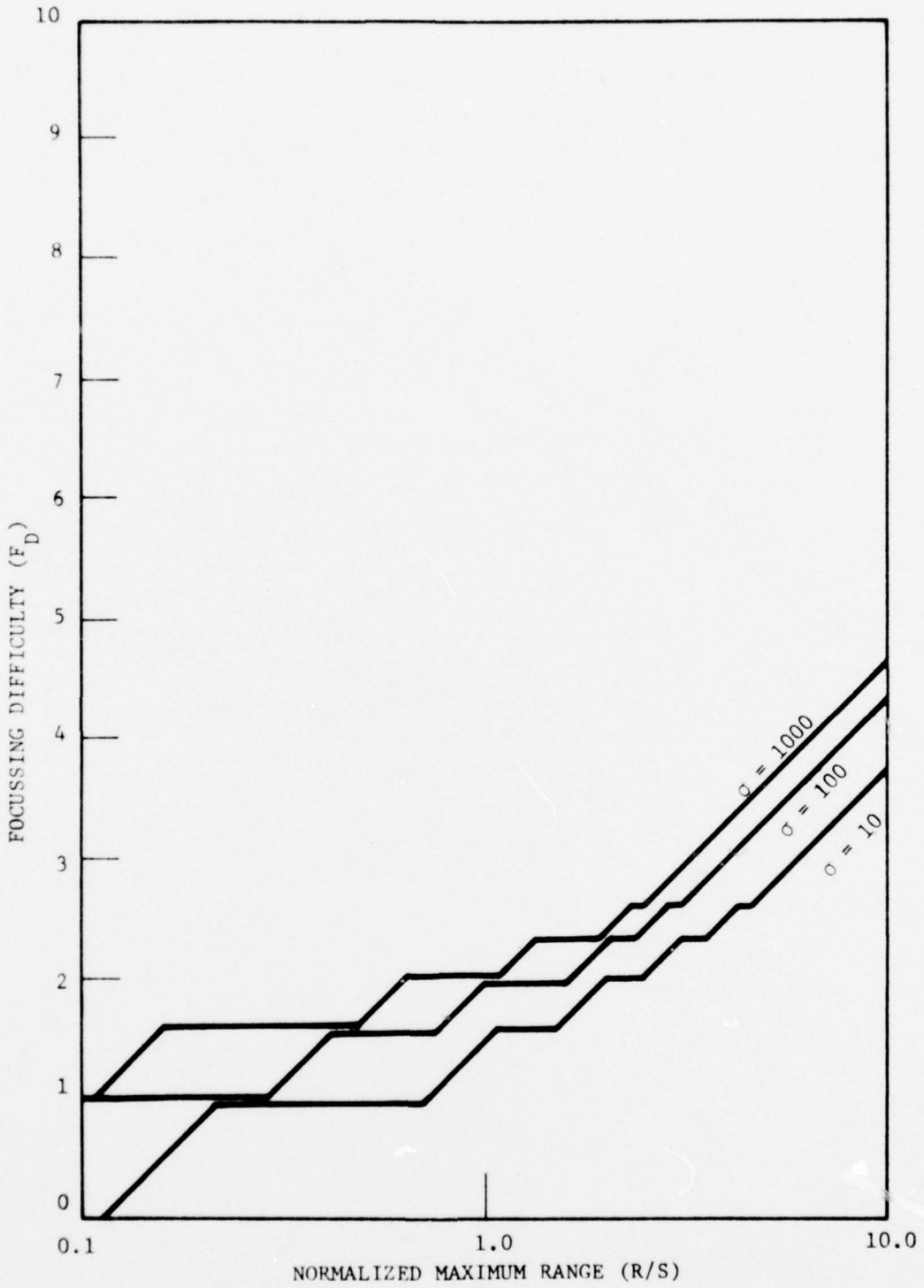


Figure 4.14 FOCUSSING DIFFICULTY USING UP TO SIX OBJECTIVES (SPATIAL RESOLUTION SPEC)

77-5-1

In some cases a system (with one or more objectives) may be designed to perform a specific target detection function up to some maximum range R. It may additionally be required that the system be kept in-focus, using the last (longest range) objective, for all ranges from R (the nominal "maximum") out to infinity. The spatial resolution will not, of course, be maintained but, subjectively, the system should appear to be sharply focused at these larger ranges.

The difficulty in performing this additional focusing task will be estimated in terms of the additional number (M') of fixed focus steps required.

From equation (4.13):

$$M' = \left(\frac{\sigma - 1}{\sigma}\right)\frac{\delta}{r}$$

$$= -\frac{\delta}{r} \text{ (for } \sigma = \infty)$$

where r is the minimum range for this additional focusing region - which is, of course, the maximum range (R) for the nominal region.

From equation (4.17):

$$\delta/R = R/S$$

$$M' = \delta/r = R/S$$

The additional focus difficulty is thus F_D'

where

$$F_D' = \log_2 (R/S)$$

The total focus difficulty, including this additional component, is shown in Figure 4.13, as a dotted line. (There is no additional focus difficulty when $R > S$.)

Application

To illustrate the application of these results, consider the problem of focusing a FLIR from 10m to 10 km for the identification of a 2m target, using a 10 μm FLIR. From Figure 4.13, it can be seen that the minimum focusing difficulty is

$$F_D \approx 4.5$$

The critical aperture (A_o) and the critical focus distance for this example are given by Table 4.1.

$$A_o = 0.31m$$

$$\delta = 2.44 \text{ km}$$

The actual aperture diameter (A) can then be derived from equation (4.22):

$$A/A_o = R/S$$

i.e.,

$$A = 0.31 \times \frac{10}{2.44}$$

$$= 1.27m$$

Figure 4.10 shows that $\rho \approx 2$ for a range-ratio of 10^3. With this value of ρ, the range-ratios and power-ratios for a three-lens system meeting the focusing specification can be read-off from the same figure (or can be easily calculated):

$$\sigma_1 = P_{12} = 1 + \rho = 3$$

$$\sigma_2 = P_{23} = 1 + \rho(1 + \rho) = 7$$

$$\sigma_3 = P_{34} = 1 + \rho(1 + \rho)(1 + \rho(1 + \rho)) = 43$$

Thus the optimum three-lens system to meet the specified focusing requirements is:

A_1 = 1.27m diameter; ranges 10 km - 3.33 km

A_2 = 0.42m diameter; ranges 3.33 km - 0.476 km

A_3 = 0.06m diameter; ranges 476m - 11m

The range ratio (σ) that can be covered by a given lens is given by

$$\sigma = 1 + MR/\delta = 1 + MS/R$$

where M is the number of fixed focus steps involved.

That is

$$M = \frac{R}{S} (\sigma - 1)$$

For the first lens:

$$\delta = R^2/S = \frac{100}{2.44} \simeq 40 \text{ km}$$

$$R = 10 \text{ km}$$

$$\sigma = 3$$

$$\therefore M = \frac{100}{2.44} \times 2 \simeq 8$$

The individual range-ratios for this lens are given by

$$\sigma_m = 1 + MR/\delta = 1 + \frac{M \times 10}{40} = 1 + M/4$$

That is

$$\sigma_1 = 1.25 \quad (10 \text{ km} - 8 \text{ km})$$

$$\sigma_2 = 1.50 \quad (8 \text{ km} - 6.67 \text{ km})$$

$$\sigma_3 = 1.75 \quad (6.67 \text{ km} - 5.71 \text{ km})$$

$$\sigma_8 = 3 \quad (3.64 \text{ km} - 3.33 \text{ km})$$

For the second lens (A2)

$$\delta = \frac{3.33^2}{2.44} = 4.5 \text{ km}$$

$$R = 3.33 \text{ km}$$

$$\sigma = 7$$

$$M = \frac{R}{S} (\sigma - 1) = \frac{3.33}{2.44} \times 6 = 8.2$$

The individual range ratios for this (second) lens are given by

$$\sigma_m = 1 + MR/\delta = 1 + m\,\frac{3.33}{4.5} \approx 1 + 0.75m$$

That is

$$\sigma_1 = 1.75 \quad (3.33 \text{ km} - 1.90 \text{ km})$$

$$\sigma_2 = 2.5 \quad (1.90 \text{ km} - 1.33 \text{ km})$$

$$\sigma_3 = 3.25 \quad (1.33 \text{ km} - 1.02 \text{ km})$$

$$\sigma_8 = 7 \quad (0.533 \text{ km} - 0.476 \text{ km})$$

For the third lens (A3)

$$\delta = \frac{(0.476)^2}{2.44} \sim 0.093 \text{ km}$$

$$R = 0.476 \text{ km}$$

$$\sigma = 43$$

$$\therefore M = \frac{R}{S}(\sigma - 1) = \frac{0.476}{2.44} \times 42 = 8.19$$

The individual range ratios for the third lens are given by

$$\sigma_m = 1 + MR/\delta = 1 + m\,\frac{0.476}{0.093} \simeq 1 + 5.25m$$

That is

$$\sigma_1 = 6.25 \quad (476m - 76.16m)$$

$$\sigma_2 = 11.5 \quad (76.16m - 41.4m)$$

$$\sigma_3 = 16.75 \quad (41.4m - 28.4m)$$

$$\sigma_8 = 43 \quad (12.6m - 11m)$$

This example has illustrated a very difficult focusing task. (Correspondingly, the system in question has a very large aperture diameter (1.27m).) The focusing difficulty, and the aperture diameter, can both be substantially reduced by limiting the maximum range (R) to 2.5 km with a range-ratio (σ) of 100, and with extension of focusing from 2.5 km to infinity at constant-angular (reducing-spatial) resolution.

From Figure 4.13, it can be seen that the focusing difficulty is now

$$F_D \approx 3$$

Equation (4.22) gives the new aperture diameter

$$A/A_o = R/S$$

i.e.,

$$A = 0.31 \times \frac{2.5}{2.44} = 0.317\text{m}$$

From Figure 4.10, the value of ρ corresponding to $\sigma = 100$ is approximately 1.25. The power ratios are then given by:

$$\sigma_1 = P_{12} = 1 + \rho = 2.25$$

$$\sigma_2 = P_{23} = 1 + \rho(1 + \rho) = 3.81$$

$$\sigma_3 = P_{34} = 1 + \rho(1 + \rho)\left(1 + \rho(1 + \rho)\right) = 11.72$$

Thus the optimum three-lens system meeting the new focusing spec is

A_1 = 0.317m diameter; ranges 2.5 km - 1.11 km

A_2 = 0.144m diameter; ranges 1.11 km - 0.291 km

A_3 = 0.040m diameter; ranges 291m - 25m

The semihyperfocal distances for the three lenses are

$$\delta_1 = \frac{2.5^2}{2.44} = 2.56 \text{ km}$$

$$\delta_2 = \frac{1.11^2}{2.44} = 505\text{m}$$

$$\delta_3 = \frac{0.291^2}{2.44} = 34.7\text{m}$$

The range ratios that can be covered by a single (fixed-focus) position for each lens, compared with the range ratios required, are

$$\sigma_1 = 1 + R_1/\delta_1 = 1.98 \text{ (2.25 required)}$$

$$\sigma_2 = 1 + R_2/\delta_2 = 3.2 \text{ (3.81 required)}$$

$$\sigma_3 = 1 + R_3/\delta_3 = 9.39 \text{ (11.7 required)}$$

The range ratios that can be achieved are almost sufficient. If two fixed-focus positions, per lens, are used the required range-ratios can be exceeded. Thus the focusing difficulty, in this example, is intermediated between

$$\log_2(3) = 1.59$$

and

$$\log_2(6) = 2.59$$

One more fixed-focus position (of the largest lens only) will extend to focusing range from 2.5 km to ∞ at constant angular resolution. Thus the total focusing difficulty lies between

$$\log_2(4) = 2$$

and

$$\log_2(7) = 2.8$$

4.6.4 Axial Image Displacement

It has been seen that an object can be kept in focus, as it moves from a maximum range R down to a minimum range R/σ, by M fixed-focus steps,

where

$$\sigma = 1 + MR/\delta$$

and δ is the semihyperfocal distance given by

$$\delta = A^2/2x\lambda$$

The magnitude (ΔU) of the corresponding axial motion of the image will now be derived.

Let U be the image distance when the object is at the maximum range (R). Then

$$1/U = 1/f - 1/R$$

where f is the focal length of the optical system.

Let U' be the image distance when the object is at the minimum range (R/σ). Then

$$1/U' = 1/f\ \sigma/R$$

i.e.,

$$\frac{U' - U}{UU'} = \frac{\sigma - 1}{R} = \frac{M}{\sigma}$$

Since

$$UU' \sim f^2$$

$$\Delta U = U' - U = \frac{Mf^2}{\delta} = \frac{Mf^2\ 2\ x\ \lambda}{A^2} = 2\ x\ \lambda\ MF^2$$

where F is the F/# of the optical system.

This equation shows that the axial image motion is the same ($2\ x\ \lambda F^2$) for each of the fixed-focus steps.

It also shows that if the system is not refocused at each fixed-focus step, the diameter of the defocus blur increases by $2\ x\ \lambda\ F$ per step.

The angular size of each pixel is α in angular units, and fα in spatial units (at the focal plane).

Using the relationship

$$\alpha = \frac{K\lambda}{A}$$

it follows that the pixel dimension is

$$\frac{K\lambda f}{A} = K\lambda F$$

Thus the increment in defocus blur per focus step (with the system not being refocused) is

$$\frac{2x\lambda F}{K\lambda F} = \frac{2x}{K} \sim 2$$

pixels per step.

REFERENCES

4.1 Thermal Imaging Systems, J.M. Lloyd, Plenum Press (pages 247-250).

4.2 Modern Optical Engineering, Warren J. Smith, McGraw-Hill Book Company (pages 318-324).

4.3 Perception of Displayed Information, J.M. Biberman, Plenum Press (pages 183-187).

SECTION 5

CONTRAST ENHANCEMENT, SHADES OF GRAY

A serious limitation of FLIR systems is the dynamic range of the display-eye interface. The FLIR video may contain scene temperature information over a dynamic range of several hundred to one (scene-temperature-range/NETD). The CRT display, however, can show only a relatively small number of distinguishably different shades of luminance. The result is that either:

(1) the CRT-gain ("contrast" control) can be reduced to avoid screen saturation at the cold and hot extremes of the FLIR image -- in which case the eye will not be able to detect small temperature differences in the scene;

(2) the CRT-gain can be set high so that the eye can detect (for example) a 0.2°C temperature change -- in which case much of the scene may saturate the display (into peak white and black), thereby making any targets in these areas totally invisible;

(3) the operator must continuously adjust his screen brightness and contrast controls performing, in effect, a "level-slice" search -- in which case the operator's workload, and also the time taken to acquire a target, are increased.

This dynamic range problem is particularly significant for rapid target acquisition in a single place aircraft. The dynamic range problem will become even more severe as FLIR sensitivity is increased in next-generation systems.

5.1 AREA ALGORITHM

One approach to this problem is that between adjacent large ($\sim 1^{o}$ on the display) areas the eye can detect luminance differentials of less than 1 percent, whereas between adjacents small ($\sim$ 1 minute of arc) areas the luminance differential must be greater than 20 percent for detection (Figure 5.1).

This suggests the following principle:

> The NETD of the FLIR should always be displayed as a just discernible luminance change.

In other words the CRT gain (percent luminance change per degree centigrade) for large areas should be much less than for small areas. The limited dynamic range of the CRT can then be made to cover a much wider range of scene temperatures.

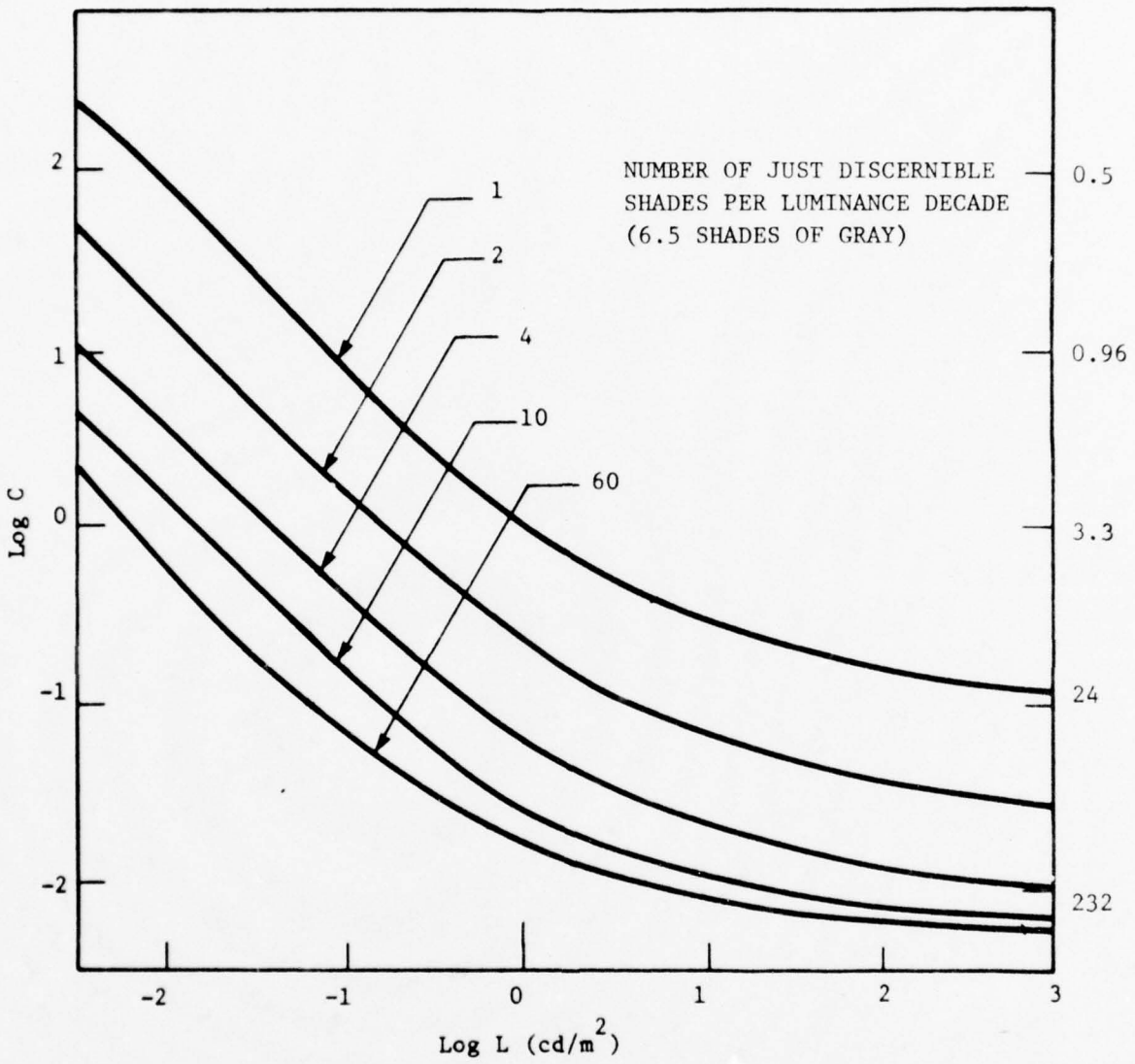


Figure 5.1 EXTRAPOLATION OF THE MEASUREMENTS OF BLACKWELL: 1-SECOND EXPOSURE OF A CIRCULAR TEST OF WHICH THE APPARENT DIAMETER (IN MINUTES OF ARC) IS NOTED ALONG EACH CURVE. THE PROBABILITY OF CORRECT ANSWERS IS 50%. ABSCISSA: DECIMAL LOG OF THE BACKGROUND LUMINANCE. ORDINATE: THE CONTRAST OF THE OBJECT COMPARED TO THE BACKGROUND. (TAKEN FROM FORM AND SPACE VISION BY YVES LeGRAND)

The FLIR image of the scene is regarded as a series of equitemperature contours T(x,y). The screen luminance function L(x,y) is chosen (for example) as

$$L(x,y) = \frac{k \Delta T\ (x,y)}{\sqrt{A(x,y)\ L(x,y)}} \qquad (5.1)$$

where L is the change in luminance between the T and T + dT equitemperature contours, A is the area of the T contour, and k is a constant.

The advantage of this area algorithm is that the large area contours, within which most of the large temperature excursions will generally fall, can be displayed at low gain and yet with the eye able to resolve the NETD. In this way, dynamic range is conserved (not wasted) (Figure 5.2). The eye is able to resolve the detail, and saturation is avoided. Within small contours, the CRT gain is higher to compensate for the lower contrast sensitivity of the eye for small areas, so that here also, the eye can resolve the NETD (Figure 5.3). However, within small contours the range of temperature excursions will generally be much less so that saturation will not occur, in spite of the high gain.

If the small contour temperature excursions should be very large, then the CRT gain can be reduced (as part of the algorithm) since in that case it will not be important for the operator to be able to resolve the NETD within the small contour.

This algorithm is, in some ways, similar to a high-frequency boost (or more accurately, low-frequency attenuation). However, there is a fundamental difference, in that the contour algorithm reduces the amplitude of large area contours without generating leading, or trailing, edge spikes (Figure 5.4).

5.1.1 Practical Implementation

A general algorithm, as in equation 5.1, may be difficult to implement particularly in real time. The following simplifications are noted, as a basis for practical implementation:

(A) Only Small Contours Need Be Considered

It is evident from Figure 5.1 that the greatest variations in contrast sensitivity fall into the lower range of contour size. This is shown in the table below:

$L = 10\ cd/m^2$

Size: Minutes of arc (~pixels)	Contrast Sensitivity (percent)
60	0.6
10	1
4	1.6
2	5
1	16

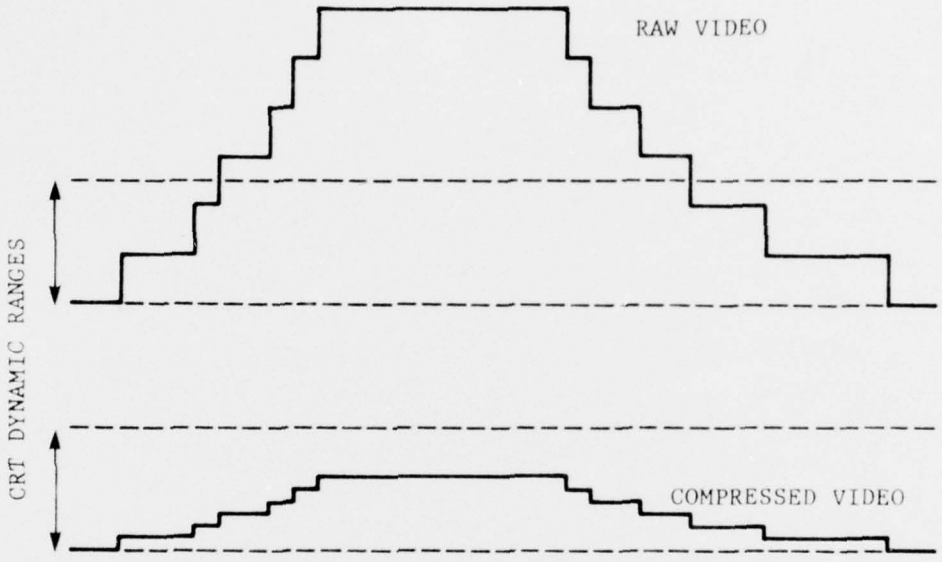


Figure 5.2 AREA-ALGORITHM COMPRESSES LARGE-AMPLITUDE SCENE DETAIL WITHOUT COMPROMISING VISUAL DETECTION

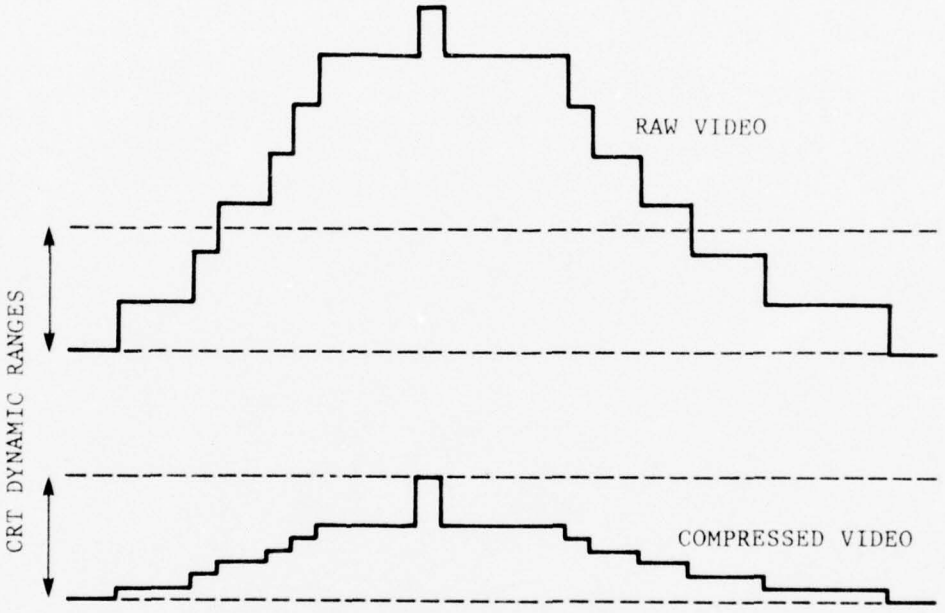


Figure 5.3 SMALL AREA DETAIL DISPLAYED AT FULL GAIN WITHOUT USING UP DYNAMIC RANGE

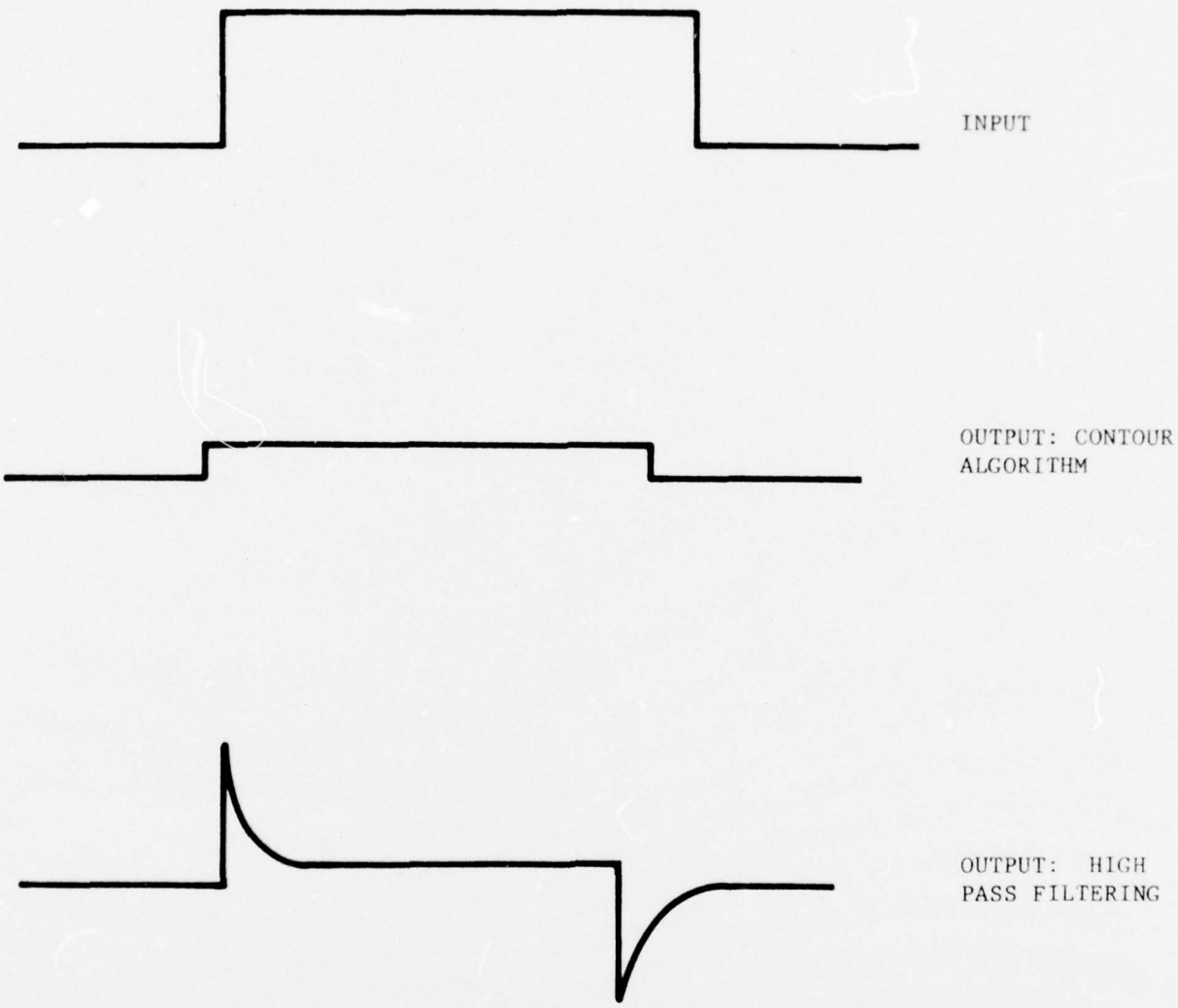


Figure 5.4 AREA-ALGORITHM COMPARED WITH A HIGH-PASS FILTER

By applying the proposed principle over contours up to 60 minutes, wide a dynamic range gain of 16/0.6 = 27 is obtained. However, if the principle is applied only over contours up to 10 minutes of arc wide, most of this dynamic range gain (16:1) is preserved. It is obviously much easier to analyze contours up to only a 10 x 10 pixel size rather than over a 60 x 60 pixel range.

In the limiting case, if the principle is applied only over 1-pixel-contours, a useful dynamic range gain of 3.2 (16/5) is obtained, and implementation becomes very simple.

(B) One Dimensional Implementation

The CRT gain adjustment can be based independently during each raster scan line, upon a one dimensional estimate of contour width. The advantage is simplicity, since there is no processing of adjacent lines, while computing the local gain adjustment. The disadvantage is that small objects will be distorted due to varying CRT gains for different parts of the same contour.

(C) Simple Extension To Two Dimensions

By considering the relative video level in the two adjacent scan lines (using storage registers, video delay lines, or special focal plane layout), it is possible to determine whether or not the current-line measure of contour width should be increased to reflect a more accurate measure of contour size then can be derived (independently) on each scan line. This correction might be sufficient to correct for the small-object distortion noted in (B) above.

(D) Focal Plane Processing

Since the proposed algorithm involves, essentially, an area operation it may be easier to implement by an appropriate distribution of detector elements and processing electronics at the focal plane.

5.2 NUMBER OF RESOLVABLE SHADES

Let C(L) be the contrast sensitivity of the eye for viewing a certain test area. That is, the test area is just discernible if displayed with luminance L + δL against a background of luminance L where

$$C(L) = \frac{\delta L}{L}$$

i.e.,

$$L + \delta L = L\,(1 + C(L))$$

Let dN be the number of discernible luminance levels between L and L + dL.

Then

$$(L + dL) = L\,(1 + C)^{dN}$$

$$\log\,(L + dL) = dN \log\,(1 + C) + \log L$$

$$\log\,(1 + \frac{dL}{L}) = dN \log\,(1 + C)$$

i.e.,

$$\frac{dL}{L} = dN \log\,(1 + C)$$

Assuming that C is constant

$$\log L'/L = N \log\,(1 + C)$$

i.e.,

$$N = \frac{\log\,(L'/L)}{\log\,(1 + C)}$$

where N is the number of resolvable luminance levels between L' and L and where C is the contrast sensitivity of the eye. Typical values of N are given in the following table:

Number of Discernible Shades per Luminance Decade

L'/L = 10

Size minutes of arc	Log_{10} C	C (percent)*	N(per decade)**
60	-2.2	0.6	384
10	-2	1	230
4	-1.8	1.6	144
2	-1.3	5	47
1	-0.6	16	16

*for L = 10 cd/m^2, taken from Figure 5.1
**A decade of luminance is equivalent to approximately 6.5 shades of grey

Values of N (per luminance decade) are shown in Figure 5.1.

5.2.1 Effect of Ambient Illumination

If the screen luminance due to ambient background is L_o then the effective contrast sensitivity (C') is decreased:

$$C' = \frac{dL}{L}$$

where

$$C = \frac{dL}{L + L_o} = \frac{dL}{L} \left(\frac{L + L_o}{L}\right)$$

is the true contrast sensitivity

i.e.,

$$C' = C \left(\frac{L}{L + L_o}\right)$$

i.e.,

$$\log C' = \log C + \log \left(\frac{L}{L + L_o}\right)$$

It is evident that the contrast sensitivity falls off very rapidly once the screen luminance falls at, or below, the background luminance L_o.

5.3 SHADES OF GREY

A shade-of-grey is defined as a $\sqrt{2}$ increment in displayed luminance (Reference 5.2). A shade-of-grey is not a just discernible contrast.

The contrast threshold of the eye (a just discernible fractional change (dL/L) in displayed luminance (L)) is a function of both the luminance and of the size of the test area. It varies from less than 1 percent to infinity.

The dynamic range of a display is the ratio of the peak luminance (L max) of the display to the luminance (L min) at which the contrast sensitivity of the eye become negligible. The ratio L max/L min is expressed as shades-of-grey:

$$\text{Shades-of-grey} = 1/2 \log_2 \left(\frac{L\ max}{L\ min}\right)$$

For example, 8 shades-of-grey corresponds to a luminance ratio of $(\sqrt{2})^8 = 16{:}1$.

L min is determined by the ambient background and by cross-coupling from the light to the dark areas of the display. It is of less practical significance than L max: a decade of improvement in L max would be much more useful than a decade improvement (reduction) in L min. This is because the contrast sensitivity and resolution of the eye both fall-off with decreasing luminance.

In night operations, L max may have to be limited to preserve dark adaptation. A color (phosphor) choice optimum for dark adaptation may not be optimum for luminous efficiency, again limiting L max.

Typical values for L max vary from 50-foot lamberts (home TV) to 1000-foot lamberts. Shades-of-grey vary from 5 (home TV) to 8-10.

5.4 APPLICATION OF OCULOMETER

A totally different approach is now described that may allow the viewer to see the full video dynamic range, without continuous adjustment of the screen controls, and without saturation of the displayed image.

An Oculometer is associated with the display to automatically indicate to the system the coordinates (X_e,Y_e) of the viewer's instantaneous point-of-fixation on the display screen.

Comparator circuits (Figure 5.5) operate on (X_e,Y_e) and on the raster voltages to define the instants during which the FLIR raster scan falls within the fixated region $(X_e \pm \delta, Y_e \pm \delta)$. The FLIR video is clamped to the average value of the video within the fixated region so that the fixated video always falls at the center of the CRT dynamic range (Figure 5.6).

When this specially clamped video is displayed at full gain, the screen will never saturate at the viewer's instantaneous point-of-fixation. However, at other points on the screen, which will be seen by the viewer's peripheral vision, saturation may occur which may cause annoying flicker.

Accordingly, the gain of the displayed video is locally reduced, as necessary, to avoid saturation of these peripheral areas.

The gain-reduction control-signal is a low resolution "video" signal readout (in synchronism with the FLIR raster) from an "image" storage medium (e.g., CCD array).

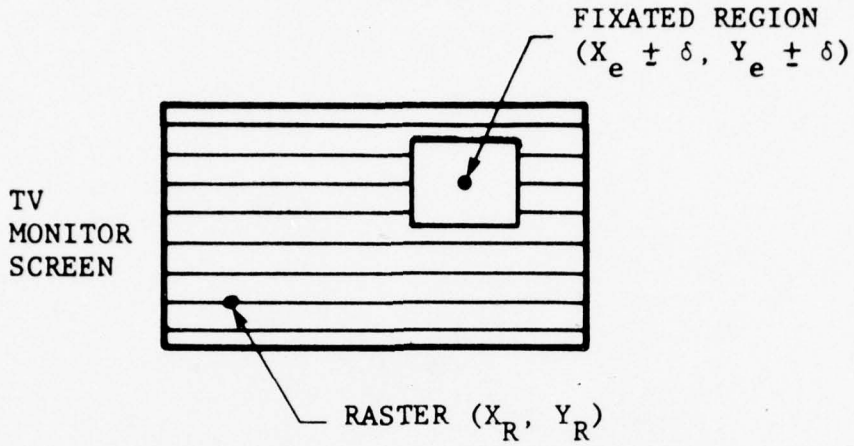
The effect to the viewer of the system (Figure 5.7) will be:

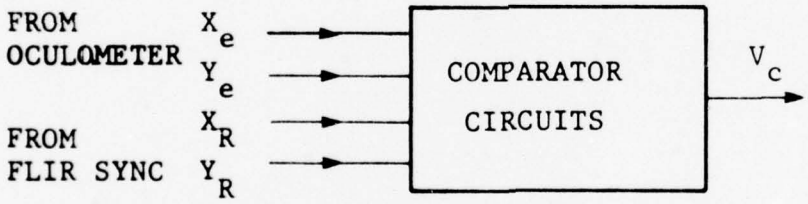
- high contrast display of the FLIR image at the point where he is looking,
- no saturation of the video anywhere on the screen,
- localized gain (i.e., contrast) reductions in very hot or very cold regions of the scene that are currently being seen by peripheral vision. Any such peripheral gain reductions should not be noticeable, since the contrast sensitivity of the eye is lower for peripheral vision than for foveal vision.

In other words, the viewer should see the full dynamic range of the video even though this might greatly exceed the dynamic range of the CRT.

A number of parameters are associated with the technique that has been described:

- The size (2δ x 2δ) of the "fixated-region."

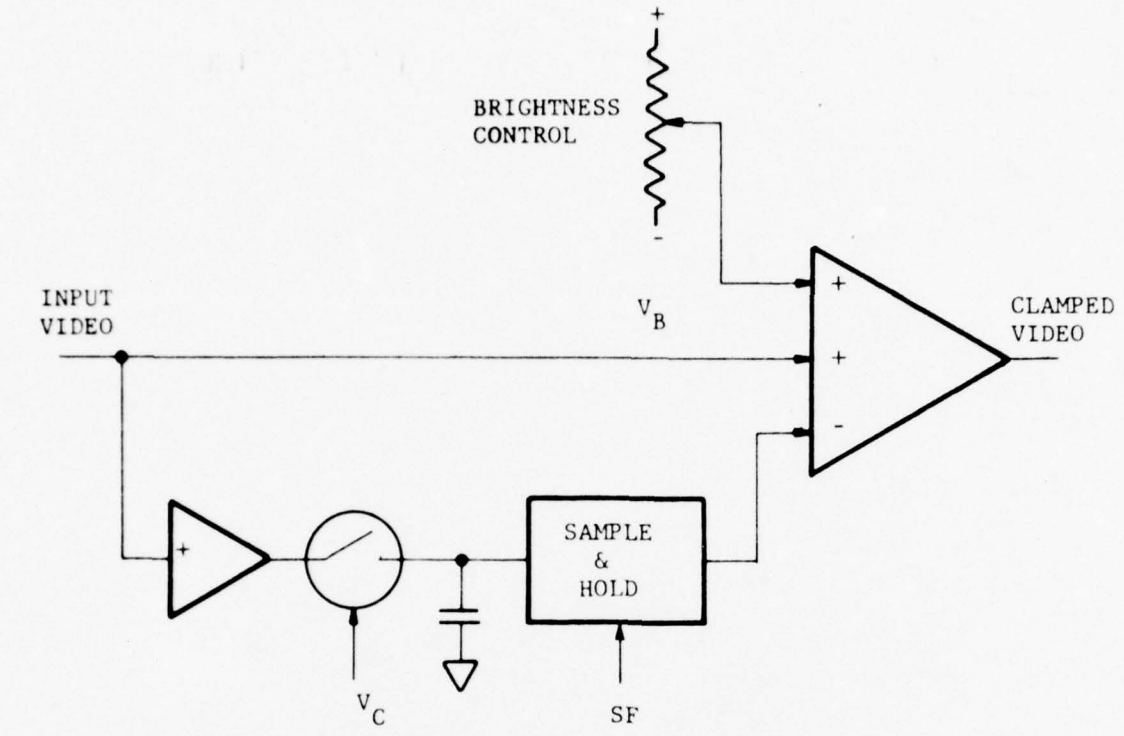




V_c = "1" $\quad (X_e - \delta < X_R < X_e + \delta$ AND
$Y_e - \delta < Y_R < Y_e + \delta)$

V_c = "0" $\quad$ OTHERWISE

Figure 5.5 DEFINITION OF FIXATED REGION



SF: START FIELD PULSE

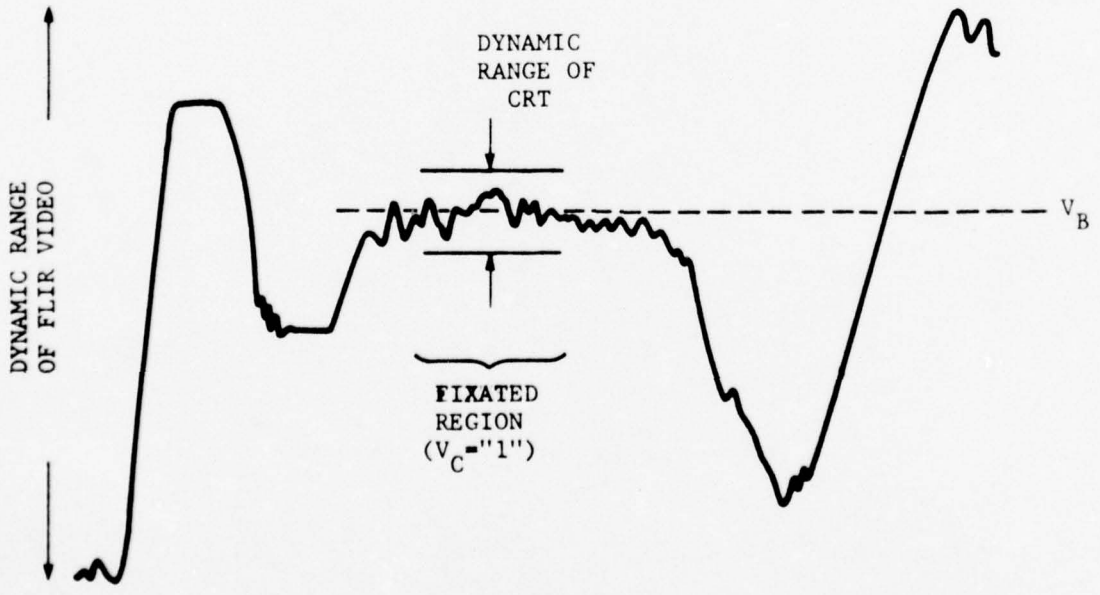


Figure 5.6 VIDEO CLAMP

77-5-1

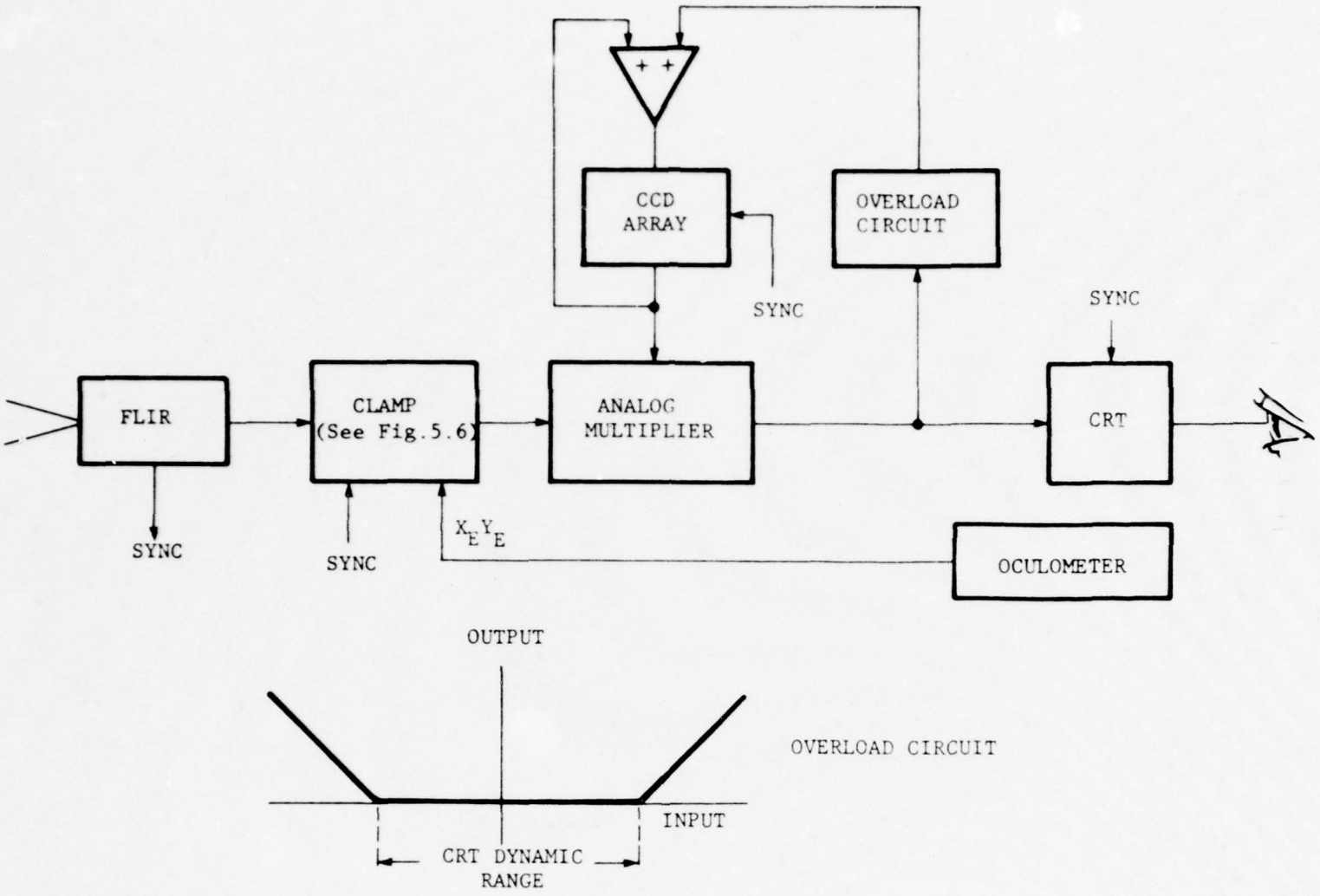


Figure 5.7 DISPLAY SYSTEM

- The resolution of the gain-controlling "image" as stored in the CCD. Relatively low resolution may be desirable -- i.e , a small hot spot in the scene should cause a smooth gain reduction over a relatively large area.
- The time-constant of the storage medium holding the gain-controlling "image." That is, the time taken for the gain.

  (a) to be reduced, in a peripheral area, to avoid saturation

  (b) to be restored to normal in a just fixated region that had previously been a saturating peripheral area.

REFERENCES

5.1 Form and Space Vision, Yves le Grand, Indiana University Press, 1967 (page 288).

5.2 Perception of Displayed Information, L.M. Biberman, Plenum Press, 1973 (page 73).

APPENDIX A

IMAGE MOTION

In Section 3 a technique is described for placing successive moving video frames in registration. This Appendix considers the limits that must be placed upon the image motion and the practical significance of such limits.

Motion Limits

If successive frames of moving video imagery are to be put into registration then certain limitations must be applied to the scene motion.

For example, the scene displacement per frame must obviously be less than one screen width, since otherwise there would be no overlapping scene area within the two successive frames. Unless there is some overlap, there is no means, by video processing of the two frames, of determining the actual displacement needed to put the frames "into registration".

These considerations set a limit to the maximum linear scene velocity in the x and y directions.

The registration sensing technique that is described in Reference 3.1 is, essentially, a null sensing technique. That is, once the address modification is found, it is used to place successive frames into registration. Successive frames, registered in this way, are then processed to detect any slight miss-registration that might exist. Any such miss-registration information is used to appropriately update the currently assumed value of the address modification. This up-dating is not a total (one-shot) correction but a change-in-the-right-direction. Thus the correction process will take several frames.

Consequently, a sudden, large, change in the frame-to-frame displacement would disturb the registration system, because the assumed address modification would then be significantly in error, and it would take several frames for this error to be corrected. This kind of disturbance is caused, not by image velocity, but by image acceleration. If the image moves at a constant velocity, the address modification is constant. Once it has been found, successive frames are always put into exact registration until some acceleration occurs.

Accordingly, a limit can be set to the unpredictable image acceleration, on the basis that the resultant registration errors (of successive frames) are kept small ($\sim$1% of screen width). In this way the blurring effects of the registration error are minimized and the error can be quickly, and unambiguously, corrected. In a similar way, limits can be set for the rotational velocity and rotational acceleration of the scene image. These motion limitations are summarized in Table A.1

(The general address modification, as derived in Reference 3.1 involving 8 scalar parameters, covers all types of image motion due to translation and rotation of the image sensor relative to the plane of the scene. For example, aspect and magnification changes, as well as simple image translations

Table A.1

IMAGE MOTION LIMITATIONS FOR REGISTRATION

Motion Parameter	Symbol	Units	Origin of Limitation	Condition	Limitation	Limitation for Sensor Operating at: 3 fps	30 fps
Image Velocity	V	Screen Width per second	Need some (conservatively 90%) overlap between successive frames in order to effect registration.	$V/R < 0.1$	$V \leq 0.1R$	0.3	3
Image Acceleration	A	Screen Width per second per second	Frame-frame registration error induced by unpredictable acceleration must be small (~1% screen width) since registration sensor is basically a null-sensor.	$A/R^2 < 0.01$	$A \leq 0.01R^2$	0.1	10
Angular Image Velocity	$\dot{\theta}$	Radians per second	Not an absolute limitation, but memory system may be simplified if all displacements are kept small	$\dot{\theta}/R < 0.1$	$\dot{\theta} < 0.1R$	0.3 ($18^\circ/s$)	3 ($180^\circ/s$)
Angular Image Acceleration	$\ddot{\theta}$	Radians per second per second	Same as for image acceleration	$\ddot{\theta}/R^2 < 0.01$	$\ddot{\theta} < 0.01R^2$	0.1 ($6^\circ/s^2$)	10 ($600^\circ/s^2$)

R = FRAME RATE OF SENSOR (FPS)

and rotation as listed in Table A.1. The rate limitations on these other parameters will be of the same order but, generally, of less practical significance).

Practical Application

In the case of a 30 fps sensor, the limitations on scene motion, shown in Table A.1., represent little or no practical restraint on operation. At these limiting rates of motion, the human operator himself would derive only minimal information from the moving image display. That is, the operator limit for scene motion is equally, or more restrictive, than that of the registration - system limit on scene motion.

For a 3 fps sensor, however, Table A.1 shows that the registration system requires relatively slow-moving scene images. It can be seen that these low rates correspond, more nearly, to that needed by the operator for effective target search and recognition. For this (and other) reasons the FLIR will generally be mounted on an inertially stabilized platform when it is to be utilized in a fast, or rapidly maneuvering, vehicle. In this way the FLIR sensor can be almost totally isolated (to the order of 1 pixel or less) from the angular (azimuth and elevation) motion of the vehicle. A platform sensor system can also be designed to stabilize against roll motion, although this feature is often not provided.

An inertially stabilized platform can isolate the sensor from x-y image displacement caused by angular, but not (by itself) from the translational, motion of the aircraft. To illustrate this, four mission scenarions are shown in Figures A.1, A.2, A.3 and A.4. For example the vertical motion of the helicopter, in the pop-up maneuver, will cause the scene image to move in the vertical direction.

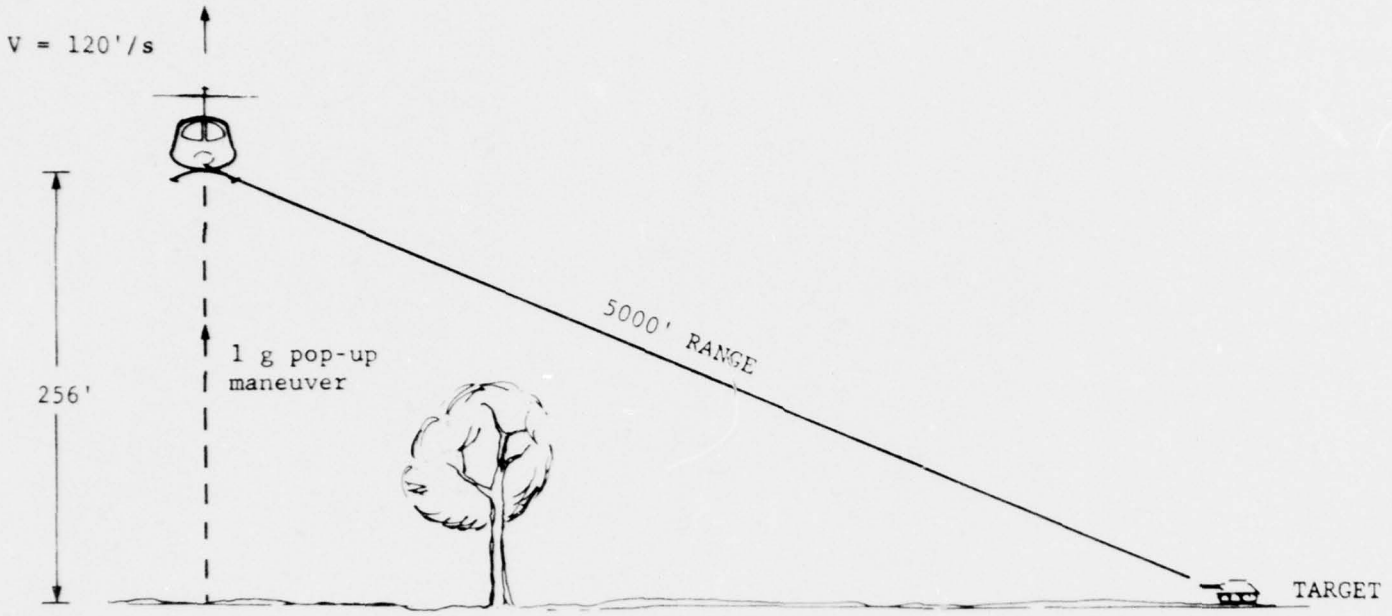
Scene image motion due to aircraft translation relative to the scene can be compensated-for if the aircraft velocity acceleration and position are known relative to the scene. Such information may be available from the aircraft navigational system, from a rangefinder, and from operator inputs of estimated target location. For the purpose of the present discussion it will be assumed that all such navigational information bears a 10% error, so that its application (to automatically steer the platform so as to compensate for aircraft motion) could reduce the displayed scene velocity and acceleration by a factor of 10.

With this assumption Table A.2 summarizes the applicability of a 3 fps sensor registration system to the four scenarios considered, with and without the use of aircraft navigational information.

Table A.2 is illustrative only, to show the possible range of application of one particular registration system application - namely the 3 fps sensor concept. It may be recalled that for a 30 fps sensor, the image motion restriction (for registration) is much less stringent. Registration, with a 30 fps sensor, is applicable in all four cases without the use of aircraft data.

IMAGE:	DISPLAYED IMAGE MOTION	
	STABILIZED PLATFORM ONLY	WITH A/C DATA
$VELOCITY_{(max)}$	24 mr/s = .24 SW/s	.024 SW/s
ACCELERATION	6 mr/s/s = .06 SW/s/s	.006 SW/s

SW = SCREEN WIDTH

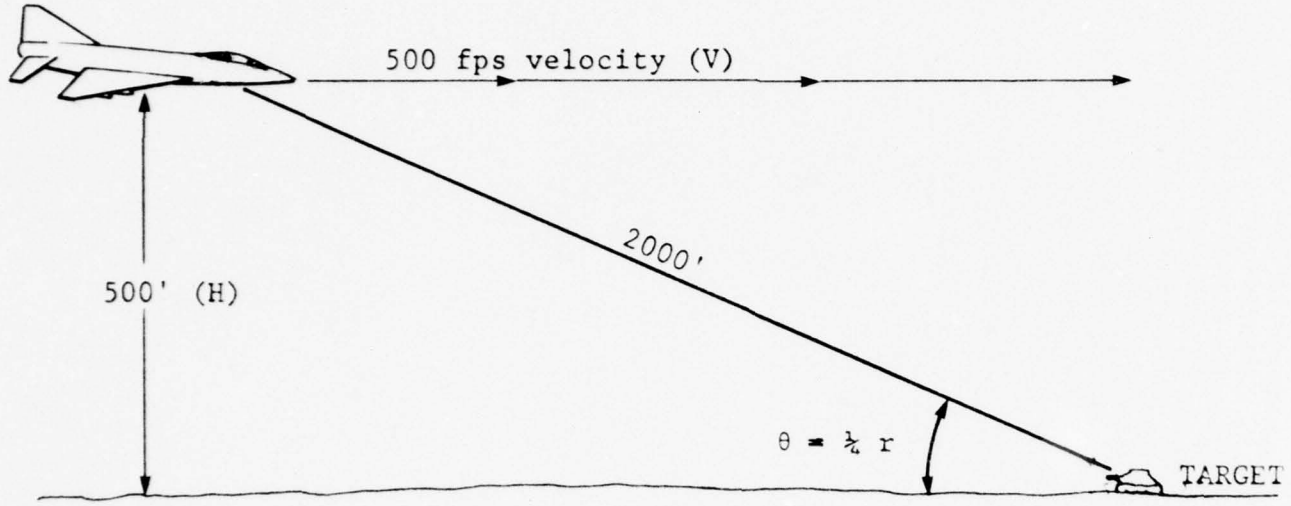


SENSOR FOV ~ 100mr

Figure A.1 1f POP-UP MANEUVER (HELICOPTER)

	DISPLAYED IMAGE MOTION	
	STABILIZED PLATFORM ONLY	WITH A/C DATA
IMAGE VELOCITY	0.6 SW/s	0.06 SW/s

SW = SCREEN WIDTH



SENSOR FOV(α) ~ 100 mr

$$\text{DISPLAYED IMAGE VELOCITY} = \frac{V \sin^2 \theta}{H \alpha} \quad (SW/s)$$

Figure A.2 LOW ALTITUDE PASS (FIXED WING)

IMAGE:	DISPLAYED IMAGE MOTION	
	STABILIZED PLATFORM ONLY	WITH A/C DATA
VELOCITY	100 mr/s = 1 SW/s	.1 SW/s
ACCELERATION	16 mr/s/s = .16 SW/s	.016 SW/s/s

SW = SCREEN WIDTH

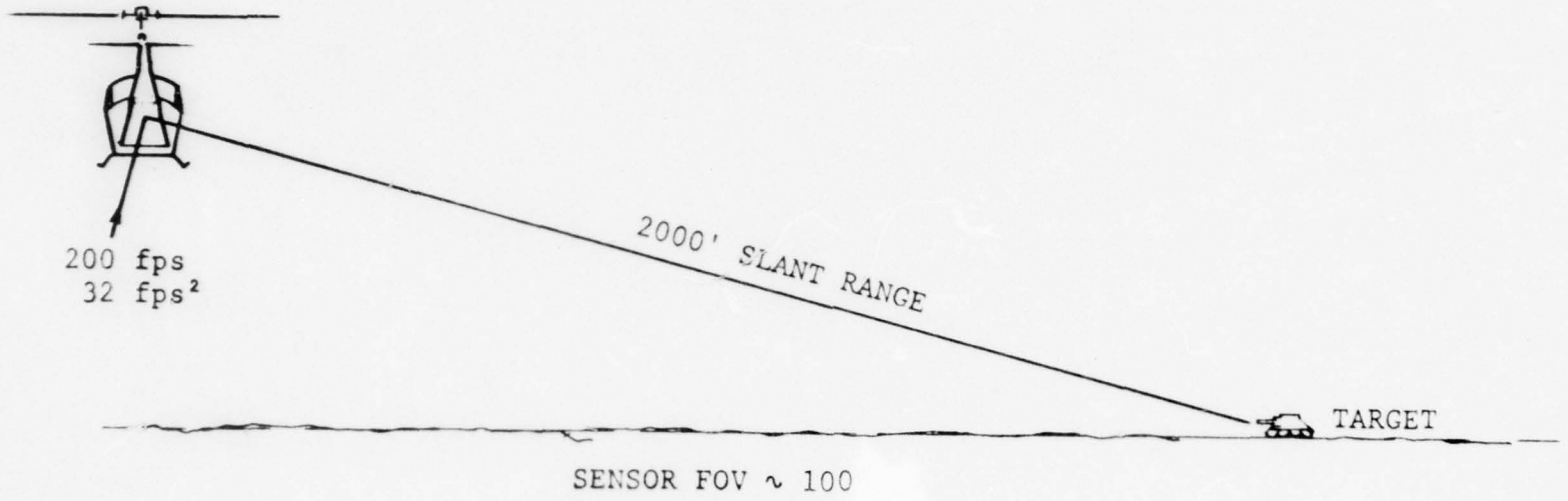
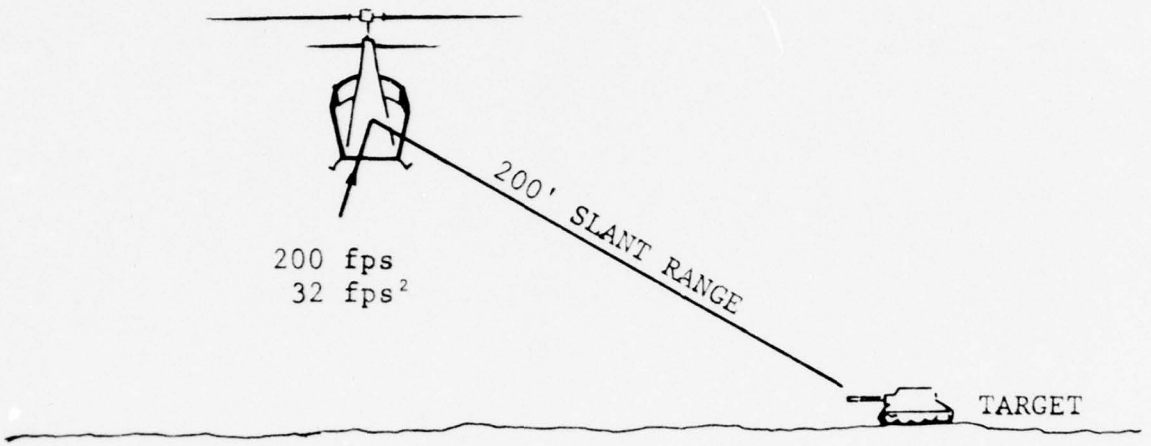


Figure A.3 CROSS FIRE: LONG RANGE (HELICOPTER)

IMAGE:	DISPLAYED IMAGE MOTION	
	STABILIZED PLATFORM ONLY	WITH A/C DATA
VELOCITY	1 r/s = 1 SW/s	.1 SW/s
ACCELERATION	.16 r/s/s = .16 SW/s/s	.016 SW/s/s

SW = SCREEN WIDTH



SENSOR FOV ~ 1 radian

Figure A.4 CROSS FIRE: SHORT RANGE (HELICOPTER)

Table A.2

IMAGE MOTION RESTRICTIONS FOR 3 FPS SENSOR:
SUMMARY OF RESULTS (ILLUSTRATIVE ONLY)

3 FPS SENSOR: REGISTRATION SYSTEM MOTION LIMIT	POP-UP		LOW ALT. PASS		CROSSFIRE: LONG RANGE		CROSSFIRE: SHORT RANGE	
		WITH A/C DATA		WITH A/C DATA		WITH A/C DATA		WITH A/C DATA
0.3 (UNITS SW/S)	0.24	0.024	0.6	0.06	1.0	0.1	1.0	0.1
0.1 (UNITS SW/S/S)	0.06	0.006	--	--	0.16	0.016	0.16	0.016
3 FPS REGISTRATION SYSTEM APPLICABLE?	YES	YES	NO	YES	NO	YES	NO	YES

APPENDIX B

EFFECT OF FRACTIONAL PIXEL ERRORS

B.1 INTEGRATION

To integrate a continuous image function V (x,y) in digital memory, it must first be sampled into a discrete array:

$$V_{ij} = \iint V(x,y)\, dx\, dy$$

where the range of integration is:

$$\frac{D}{M}(M-i) \leq x \leq \frac{D}{M}(M+1-i)$$

$$\frac{E}{N}(N-j) \leq y \leq \frac{E}{N}(N+1-j)$$

where

$$1 \leq i \leq M$$

$$1 \leq j \leq N$$

This corresponds to a discrete representation of V over an M x N pixel array with screen dimensions D x E.

Consider the particular case of a sinusoidal image, given by

$$V(x,y) = A \sin \frac{2\pi f}{D}(x+\xi)$$

where ξ is an x displacement that will vary from frame-to-frame (That is, ξ will represent the motion of the image). This sinusoid is at f cycles per frame width.

Let V_{ij}^{1}, V_{ij}^{2}, V_{ij}^{3}..... be the discrete arrays corresponding to ξ values ξ_1, ξ_2, ξ_3.....

If the displacements ξ_1, ξ_2, ξ_3 ....etc correspond to exact integral charges in i, that is

$$\Delta\xi \;=\; \text{INTEGRAL MULTIPLE OF } \frac{D}{M} \;=\; \Delta i$$

(where Δi is an integer) then the various arrays $V_{ij}^{\,1}$, $V_{ij}^{\,2}$, $V_{ij}^{\,3}$, etc can be put into <u>exact</u> registration by an integral shift (address modification) in i.

In general, however, the change $\Delta\xi$ in ξ between frames will <u>not</u> correspond to an integral-value change in i. Then the arrays $(V_{ij}^{\,K})$, can never be put into exact registration.

By choosing the closest integral value change in i between the arrays (corresponding to the actual displacement $\Delta\xi$) the best possible registration of the arrays is obtained.

The image degradation resulting from this compromise (i.e. inexact) registration will now be derived.

The discrete sample V_{ij} is given by:

$$V_{ij} = \iint A \sin \frac{2\pi f}{D} (x+\xi)\, dx\, dy$$

$$= \frac{DA}{2\pi f} \left[\cos 2\pi \frac{f\,(x+\xi)}{D} \right] \left[y \right]$$

where the range of integration is

$$\frac{D}{M}\,(M-i) \leq x \leq \frac{D}{M}\,(M+1-i)$$

$$\frac{E}{N}\,(N-j) \leq y \leq \frac{E}{N}\,(N+1-j)$$

Thus

$$V_{ij} = \frac{DA}{2\pi f} \frac{E}{N} \left\{ \cos \frac{2\pi f}{D} \left[\frac{D}{M} (M+1-i) + \xi \right] - \cos \frac{2\pi f}{D} \left[\frac{D}{M} (M-i) + \xi \right] \right\}$$

$$= \frac{DE}{2\pi f N} 2A \sin \left[\frac{2\pi f}{M} (M-1+\tfrac{1}{2}) + \frac{2\pi \xi f}{D} \right] \sin \frac{\pi f}{M}$$

As ξ changes, integral changes in i can be made to approximately compensate (i.e., to match the x shift). However, there will, in each case, be a residual component of the angle $\frac{2\pi \xi}{D}$ that cannot be compensated for by an i-shift. This residual angular error will, in general, be uniformly (randomly) distributed and can be represented by

$$\theta_K = \frac{\pi f}{M} \alpha (K)$$

where $-1 \leq \alpha (K) \leq +1$

Then the sum (i.e., average) of the approximately registered arrays is

$$V_{ij} = \sum_k V_{ij}^k = \sum \frac{DEA}{\pi f N} \sin \left\{ \frac{2\pi f}{M} (M-i+\tfrac{1}{2}) + \theta_k \right\} \sin \frac{\pi f}{M}$$

The mean value of V_{ij} (over the random distribution of θ_K) is given by:

$$\overline{V} = \frac{DEA}{\pi f N} \sin \frac{\pi f}{M} \int_{\frac{-\pi f}{M}}^{\frac{+\pi f}{M}} \sin \left[\frac{2\pi f}{M} (M-i+\tfrac{1}{2}) + \theta \right] \frac{d\theta}{\left(\frac{2\pi f}{M} \right)}$$

$$= \frac{DEA}{\pi f N} \sin \frac{\pi f}{M} \quad \frac{\left[\cos \frac{2\pi f}{M} (M-i+\frac{1}{2}) + \theta\right]_{\frac{-\pi f}{M}}^{\frac{+\pi f}{M}}}{\frac{2\pi f}{M}}$$

$$= \frac{DEA}{\pi f N} \sin \frac{\pi f}{M} \quad \frac{\sin \frac{2\pi f}{M} (M-i+\frac{1}{2}) \quad \sin \frac{\pi f}{M}}{\frac{\pi f}{M}}$$

The attenuation of the average $(\bar{V})$ relative to a perfectly registered single frame is

$$V_{Rel} = \frac{\bar{V}}{V_{ij}^{K} \ (\theta_K = 0)} = \frac{\sin \pi f/M}{\pi f/M}$$

This result shows that the mean amplitude of the sinuoid, over the set of sinusoids with random fractional pixel errors, is as given in Figure B.1.

The aliasing frequency f_m is given by:

$$f_m/M = 0.5$$

i.e.,

$$f_m = M/2 \quad \text{(cycles/screen width)}$$

The aliasing frequency (f_m) is the maximum spatial frequency for which the discrete samples (M/screen width) can be used to accurately recreate the original input function. If the input function contains spatial frequencies in excess of M/2 (cycles/screen width), the sampling process generates a fold-over of the spatial frequency spectrum which causes errors in the recreation of the origianl function from the samples.

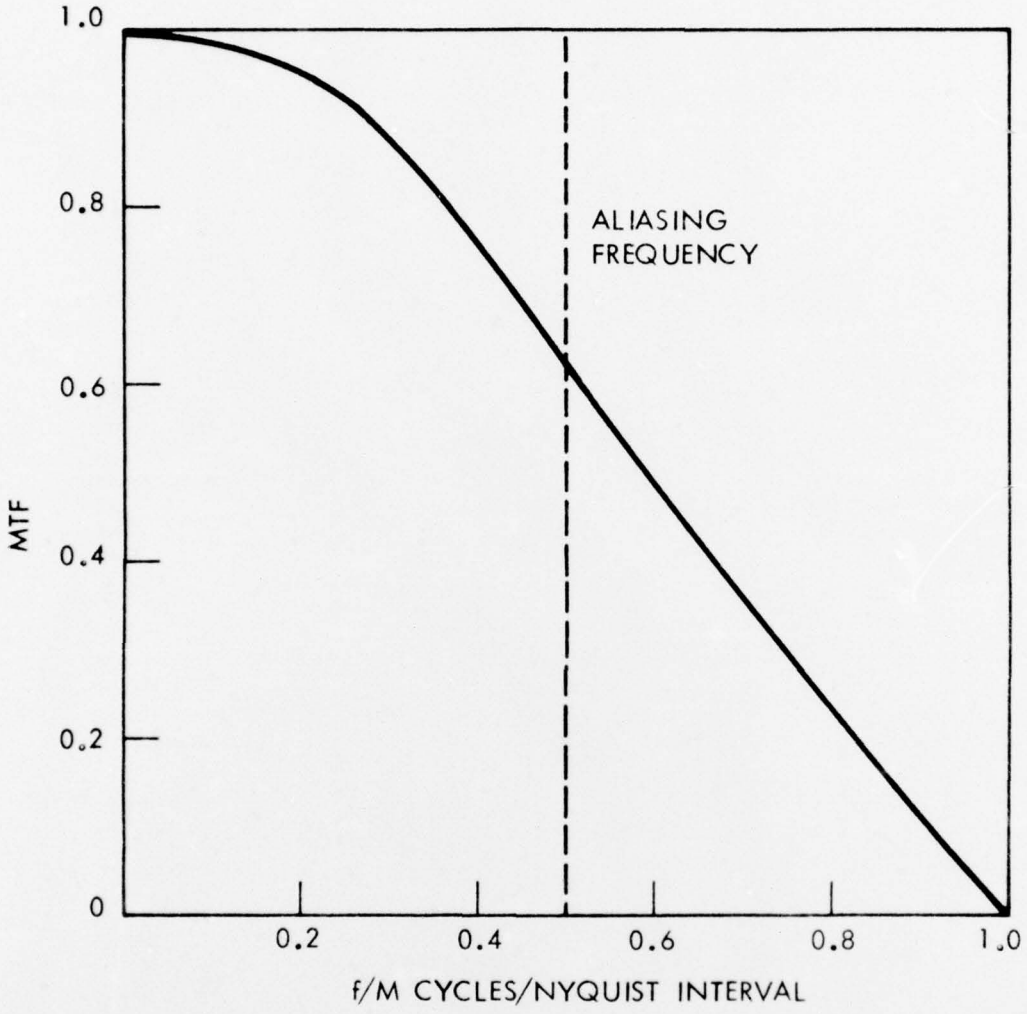


Figure B.1 MEAN MTF DEGRADATION DUE TO INTEGRATION OF FRACTIONAL PIXEL ERRORS

For this reason, the sampling frequency (M) is usually chosen to be twice that of the highest spatial frequency in the input image. Under these conditions the attenuation introduced by fractional pixel errors is minimal (Figure B1). To further illustrate this point, the example given in Table 3.5 of Reference B1 is cited. This gives the MTF's for a typical FLIR system in which the detector cut-off frequency (1 cy/m rad) is one-half that of the diffraction-limited spatial frequency (2 cy/m rad). The overall transfer function of this illustrative system (including optics, detector, electronics, and monitor), is plotted in Figure B.2, together with the attenuation introduced by integration with random fractional pixel errors.

It can be seen that the overall effect of the fractional pixel errors is negligible.

The same kind of analysis (as that presented above) can be performed in the spatial domain. Consider a band-limited, continuous-input impulse-function V(x,y) given by:

$$V(x,y) = \frac{\sin \frac{\pi Mx}{D}}{\frac{\pi Mx}{D}}$$

(M Nyquist samples/screen width D: upper cut-off spatial frequency M/2 cycles/screen width).

Because of the fractional pixel problem this function will smear out when integrated many times. The mean values of the integrated function is:

$$\bar{f} = \text{Expected Value } (f^K)$$

$$= \text{Expected Value } \left\{ \frac{\sin \pi \left(\frac{Mx}{E} + \alpha \right)}{\pi \left(\frac{Mx}{E} + \alpha \right).} \right\}$$

where α is distributed uniformly over the range $\pm \frac{1}{2}$.

$$\therefore \bar{f} = \int_{-1/2}^{+1/2} \frac{\sin \pi \left\{ \frac{Mx}{E} + \alpha \right\}}{\pi \left\{ \frac{Mx}{E} + \alpha \right\}} \, d\alpha.$$

$$= \frac{1}{\pi} \left[\text{Si} \left(\pi \left[\frac{Mx}{E} + \alpha \right] \right) \right]_{-\frac{1}{2}}^{+\frac{1}{2}}$$

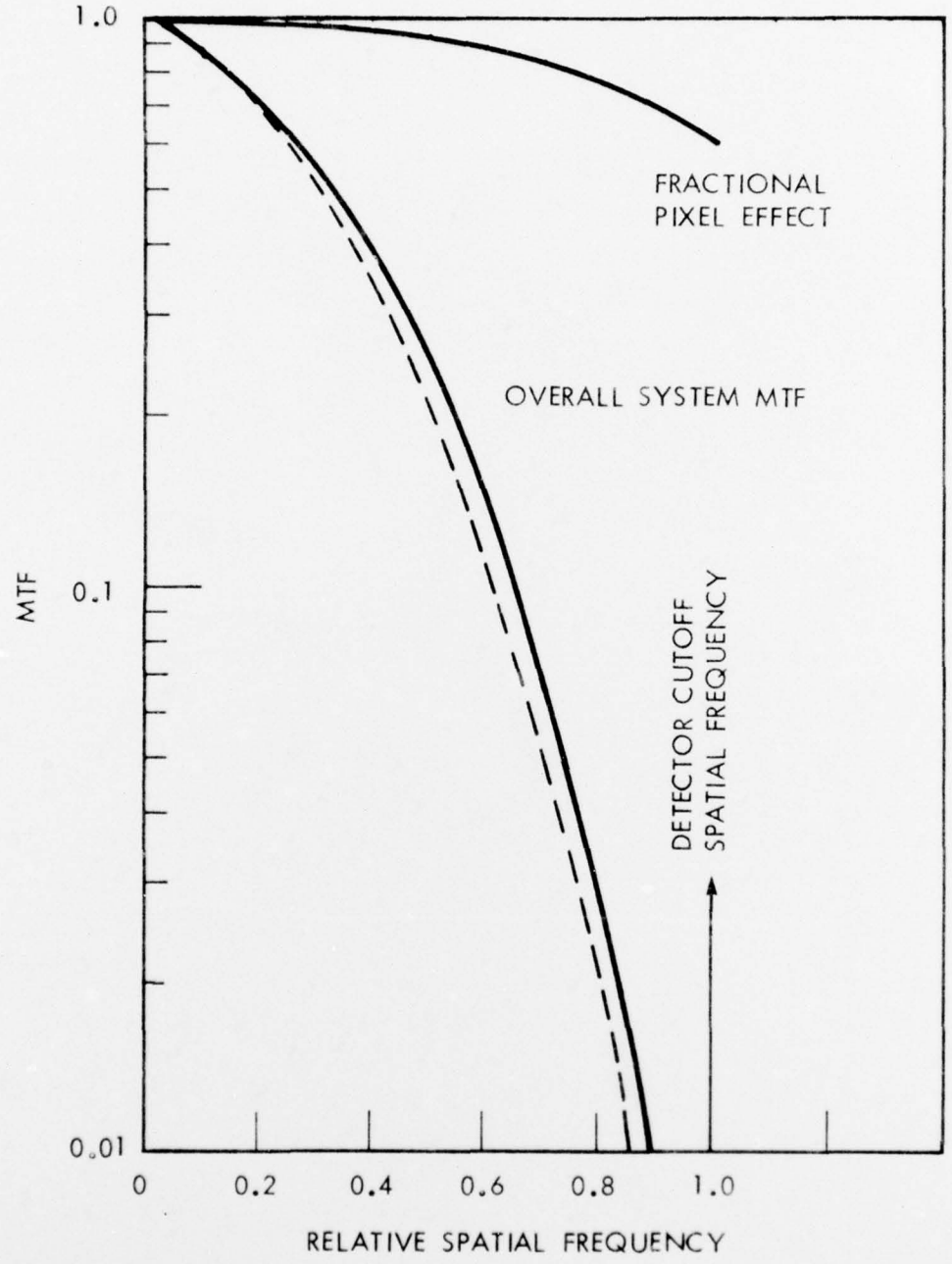


Figure B.2 EFFECT OF FRACTIONAL PIXEL ERRORS ON TYPICAL FLIR OVERALL SYSTEM MTF

where Si(x) is the sine integral:

$$Si(x) = \int_0^x \frac{Sin(u)}{u}\, du.$$

and its value is given in standard mathematical tables.

The function $\bar{f}(x)$ is plotted in Figure B.3. It can be seen that the effect of averaging the impulse function over a set of random fractional pixel errors is only a slight reduction in amplitude (consistant with the MTF degradation indicated in Figure B.1).

It is concluded, from these frequency and spatial domain results, that the effect of fractional pixel errors is negligible for the case of integration, in best registration, of a number of successive (moving) video frames.

In a typical integration application the averaging of the fractional pixel errors will be over a relatively large number of frames (e.g., 60 frames), so that the deviation from the mean value will be small.

B.2 GENERATION OF MOVEMENT

The sampling theorem shows that a band limited function can be sampled, and then exactly reconstructed, from these samples without any error at all, irrespective of any fraction pixel positioning of the input. That is, if the input function displaces by only a fraction of a pixel, then the output (reconstructed) function will also (automatically) be displaced by that same fraction of a pixel.

A fractional pixel limitation arises, however, when it is desired to shift an already sampled image, by repositioning the sample values within the sample matrix.

It has been seen that these fractional pixel errors average out to a negligible value when a number of sampled image-arrays (each with its own random fractional pixel error) are superimposed. However, when a single sampled-image-array is to be moved (by pixel repositioning) then fractional pixel errors will be significant.

For example, in one application of the registration technology, the sensor frame rate is reduced (e.g., to 3 fps) and the electronic system generates the missing 27 fps by suitably moving one sampled frame in order to simulate the effect of a 30 fps sensor viewing the moving scene.

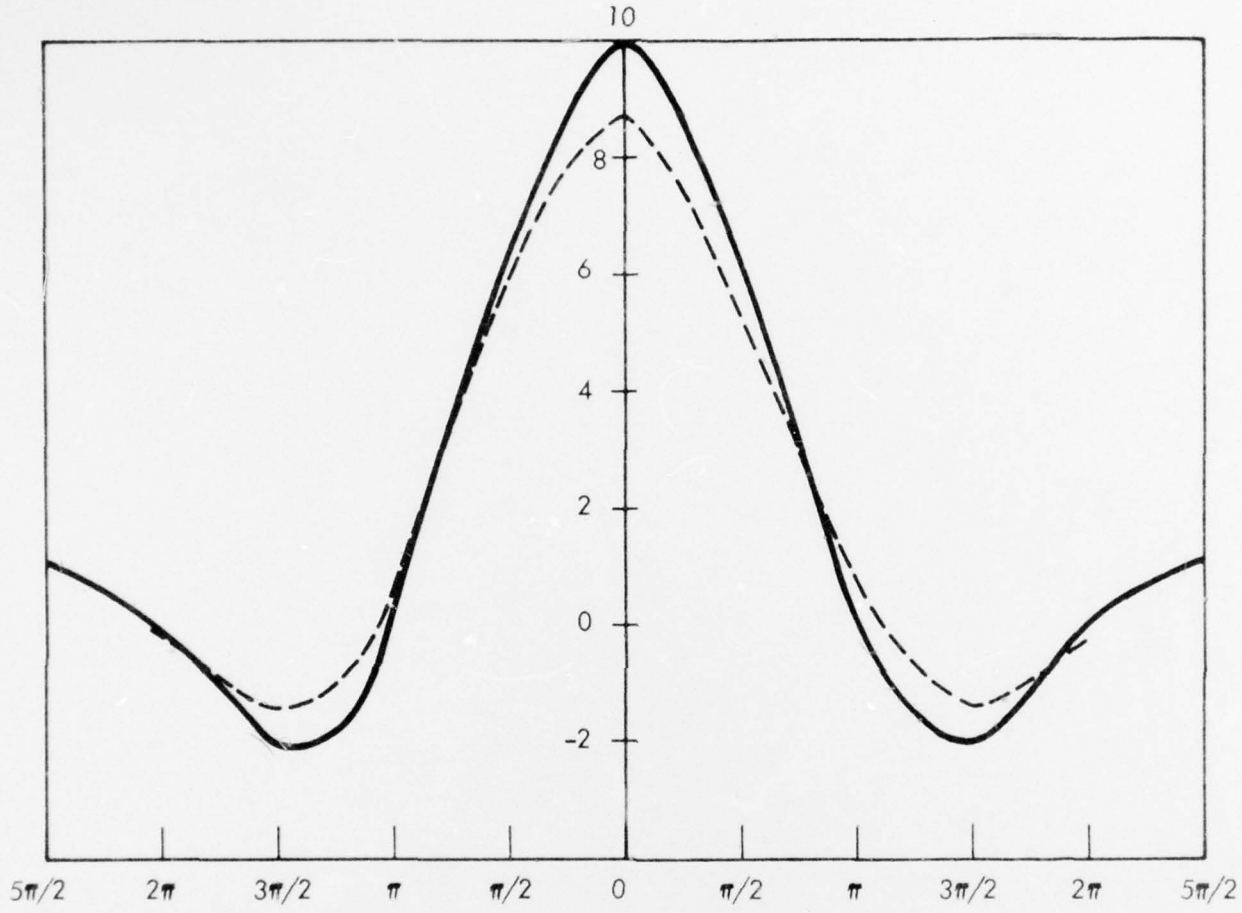


Figure B.3 DEGRADATION IN IMPULSE RESPONSE DUE TO INTEGRATION OF FRACTIONAL PIXEL ERRORS

In another application, a sampled-image-array (stored in memory) is to be artificially zoomed. That is, every 1/30 of a second, the same image is read-out from memory with an ever increasing magnification-type of address-modification.

In these applications the fractional pixel error will produce a significant effect. For example, suppose the scene consists of a horizontal line, 1 pixel wide, running exactly along a scan line. The image, in memory would be as in Figure B.4(a). Now if the image is rotated (for example) by software manipulations its representation in memory becomes as in Figure B.4(b).

This problem (illustrated in Figure B.4) arises because of failure to implement fractional pixel shifts. However, such shifts can be implemented.

For example the time series a_n can be shifted by 1/2 pixel (without distortion) by convolution with the time series b_n

$$b_n = \frac{\sin \pi (n+1/2)}{\pi (n+1/2)}$$

By truncating the series b_n a simpler implementation is obtained, but with less fidelity. For example with just two terms;

$$b_1 = b_2 = 2/\pi,$$

the transfer function is as shown in Figure B.5. By going to four terms;

$$-\frac{2}{3\pi}, \frac{2}{\pi}, \frac{2}{\pi}, -\frac{2}{3\pi},$$

the transfer function is improved, as illustrated in Figure B.6.

In the case of a two dimensional image, these convolutions must be applied in two dimensions.

The simple two-term correction (Figure B.5) would be most easily implemented by arranging for the pixel selection between two adjacent pixels (whose addresses bracket the fractionally indicated-address) to be a statistical

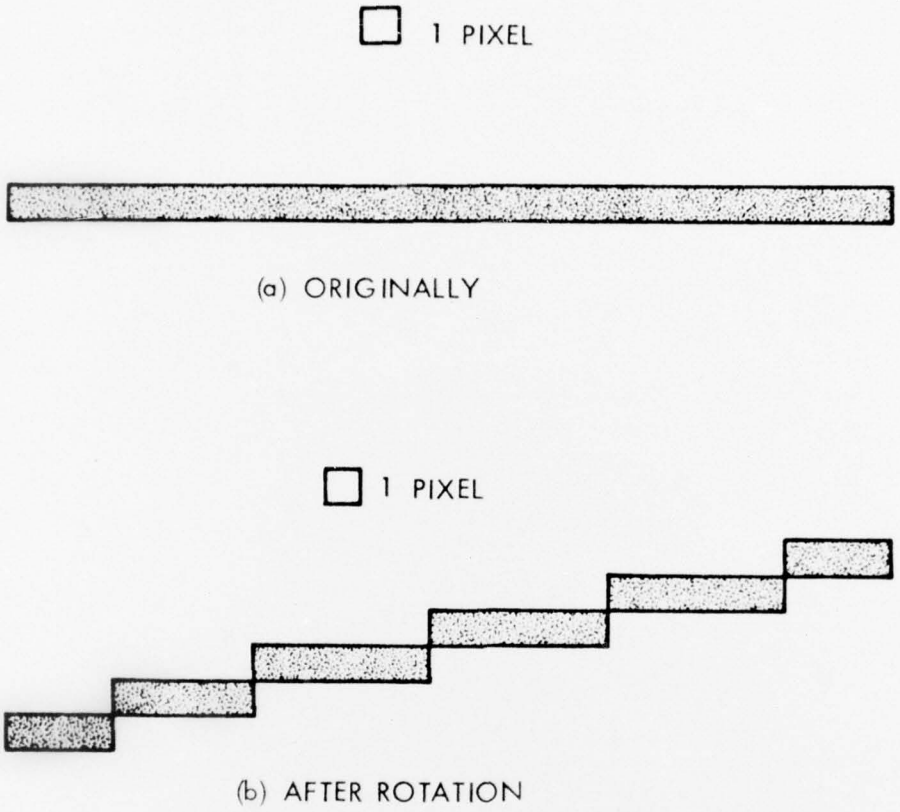


Figure B.4 OUTPUT REPRESENTATION OF STRAIGHT LINE

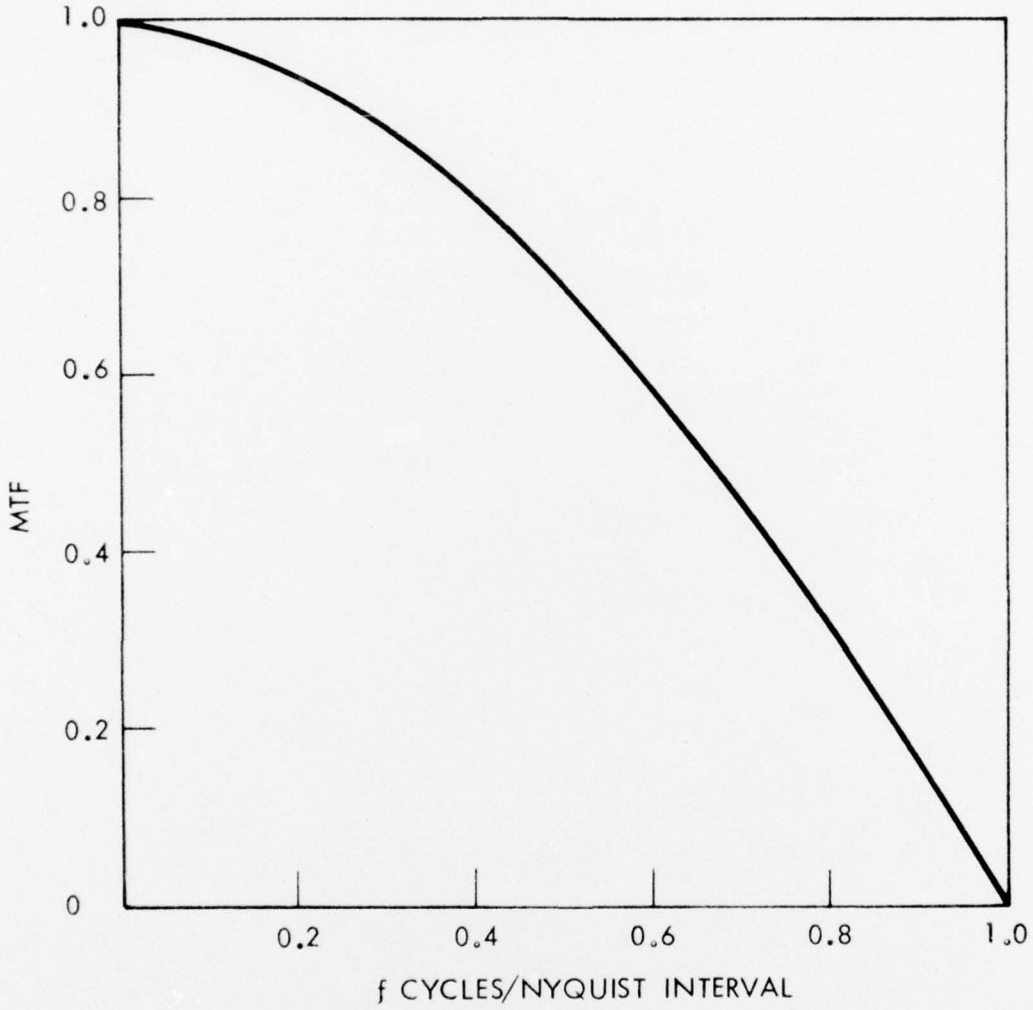


Figure B.5 MTF DEGRADATION CAUSED BY USING $(1 + e^{iwT})$ TO IMPLEMENT A ONE-HALF PIXEL SHIFT

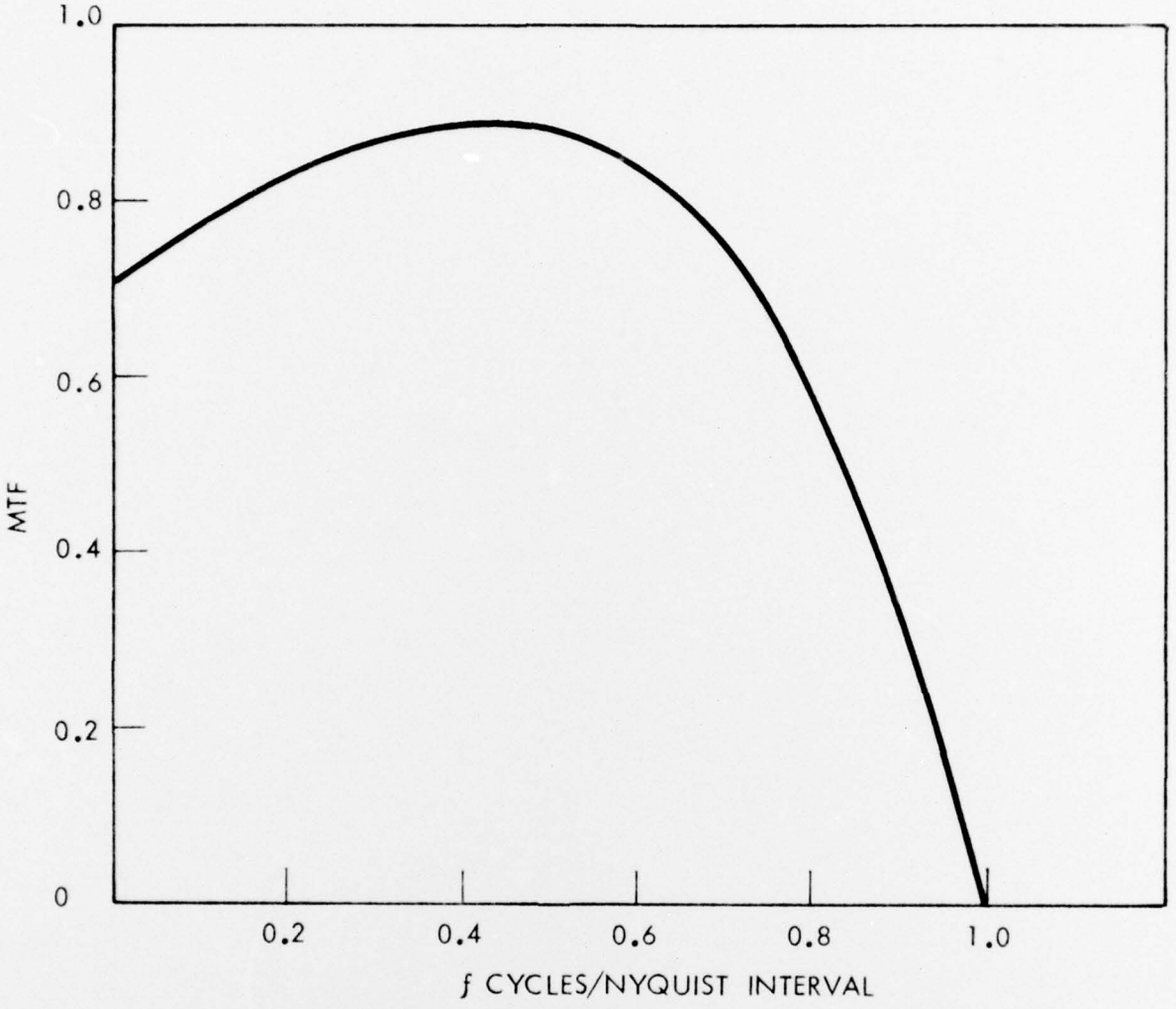


Figure B.6 MTF DEGRADATION CAUSED BY USING $(-\frac{e}{3}^{-iwT} + 1 + e^{iwT} - \frac{e}{e}^{2iwT})$ TO IMPLEMENT A ONE-HALF PIXEL SHIFT

function in both the x and y directions of the fractional component of the pixel address. For example:

fractional address 3.5:	actual addresses of 3 and 4 would be equally likely.
fractional address 3.1:	address 3 would be 10 times more probable than 4.

Multi-term corrections would be implemented in hardware utilizing line storage registers and high speed digital multipliers.

A completely different approach to the problem is to arrange for the memory-read-out system to generate a fractional pixel displacement command (in x and y) that would constitute two additional inputs into the display monitor. These two monitor inputs would be used to displace the flying spot by up to $\pm$ 1/2 pixel. In this way the image in digital memory is displayed without any fractional pixel errors.

This latter technique is simpler, in concept, but requires a specially modified monitor.

REFERENCES

B1. Thermal Imaging Systems, J. M. Lloyd, Plenum Press 1975 (p 112).